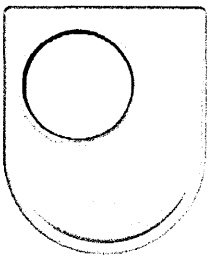
A molecular genetic characterisation of glutathione synthetase in Drosophila melanogaster

Thesis



The Open University

A molecular genetic characterisation of glutathione synthetase in *Drosophila melanogaster*

Joanne C. Daniels, B.Sc. (Hons), M.Sc.

A thesis submitted in partial satisfaction for the degree of Doctor of Philosophy

Submitted October, 2006

Supervised by Dr. R.D.C Saunders
Department of Biological Sciences
The Open University
Walton Hall
Milton Keynes

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Declaration

The work contained in this thesis is entirely my own, and is the result of my own academic and experimental enquiry. Contributions to the work by colleagues are fully acknowledged in the text. This work has not been, and is not currently being submitted for candidature for any other degree.

Acknowledgements

Completing this thesis would have been impossible without the help and support of many people over the past four years. I am especially grateful to my supervisor, Dr Robert Saunders, who patiently taught me about fly genetics, and whose supervision approach taught me to think for myself, as well as to learn from my mistakes and try out new ideas.

My lab mates were a great source of support with their practical advice about experiments, and their encouragement throughout my three years in the lab. Dr Pushpa Kansagra walked me through the project in the first few months, always willing to give her time and expertise. Dr Kansagra created the following fly stocks which were used in this thesis: *l(1)8.3/FM6*, pUAST-*Gclc* and pUAST-*Gclm*; and confirmed the position of the EP element in stock *w sn EP1322* by sequencing the insertion region. Dr Ruth Akhtar was always at hand to discuss molecular and cell culture techniques; and Claire Kotecki was patient and kind when teaching me about fly pushing. Claire Kotecki was also responsible for the remobilisation and balancing of the transgenic pUAST-*Gclc* and pUAST-*Gclm* fly stocks. I am also grateful for the insightful discussions with Dr Mark Hirst and many other of my colleagues.

The friendship and support of my fellow PhD students, Maria Tynan, Howard Donohue, Glenn Dallerac and Sarah Anderson, was invaluable - without our crazy office and chats over coffees (or wine) I don't think I would have got through it all. My South African university friends, Helen Dampier and Yvette Morey, have been helpful with my writing up, and I am grateful for their understanding over this past year, during which I haven't been able to socialise as much as I would have liked to.

Finally, my family has always provided unconditional love and support. My parents have always encouraged me in whatever I have wanted to do, and have been there for me on the other end of the phone whenever I have needed them. The biggest thanks goes to my husband, Reuben, who has been with me from the beginning of this PhD, and has had to endure all the ups and frequent downs, providing constant comfort and reassurance, and always having total faith in me.

Abstract

Glutathione synthetase (GS) catalyses the final step of glutathione (GSH) synthesis. GSH is an important antioxidant, involved in detoxifying free radicals and maintaining the cellular redox balance. Oxidative stress is thought to be one of the causes of cellular senescence and organismal aging. The broad scope of this work is the analysis of oxidative stress resistance in *Drosophila melanogaster*, in relation to ageing, and specifically, this project aims to characterise GS in *D. melanogaster*.

The gene encoding GS in *D. melanogaster* (*DmGS*) is 8Kb long, and consists of up to 10 exons. Nine distinct potential transcripts, which differ due to alternative splicing, have been identified. In this study, cDNAs obtained from adult flies and S2 cells were sequenced, and confirmed the presence of six transcripts *in vivo*. Functional complementation in yeast corroborated the bioinformatic approach used to identify the two transcripts which are likely to produce an active GS protein.

A tandem duplication of GS in certain strains of *D. melanogaster* was identified using PCR techniques and southern blotting. Analysis of a small subset of sibling species indicated the duplication was not present in these closely related sibling species, or in any wild type *D. melanogaster* strains tested.

The effect of GS overexpression on oxidative stress resistance in adult flies was investigated, and the overall contribution of the protein to GSH production in relation to GCLC was also analysed. Increased levels of transcription of both genes raised GSH levels in adult flies, and also increased oxidative stress resistance.

After attempting to create a GS null mutant fly, GS knockdown in *Drosophila* S2 cells was studied using RNAi. GS RNAi was not found to have an effect on GSH levels, although in combination with GCLC knockdown, it did have a cumulative effect on lowering cellular GSH concentrations. Knockdown of either GSH synthesis genes did not have an effect on cellular viability, although upon exposure to DEM, GS and GCLC RNAi both significantly decreased cellular survival.

Publications

Conference abstracts

Gilfillan JC, Kansangra P, Saunders RDC (2005) Glutathione synthetase and oxidative stress in *Drosophila melanogaster*. 46th Annual *Drosophila* Research Conference, San Diego.

Gilfillan JC, Kotecki C, Kansagra P, Saunders RDC (2004) Glutathione, oxidative stress and aging in *Drosophila*. Gordon Research Conference Biology of Aging meeting, Aussois France.

Contents

List of Figures

List of Tables

Abbreviations

AA	Amino acid
Amp	Ampicillin
AP-1	Activator protein-1
APS	Ammonium persulphate
ARE	Antioxidant response element
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
bp	Base pair
BSA	Bovine serum albumin
CIP	Calf intestinal alkaline phosphatase
CR	Caloric restriction
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
cDNA	Complementary DNA
DDT	Dithiothreitol
DEM	Diethyl maleate
DEPC	Diethyl pyrocarbonate
DMSO	Dimethyl sulfoxide
dGTP	Deoxyguanosine triphosphate
dH_2O	Distilled water
Da	Dalton
DmGS	*Drosophila melanogaster* glutathione synthetase
dNTP	Deoxynucleotide triphosphate
DNA	Deoxyribonucleic acid
DR	Dietary restriction

dsRNA	Double stranded RNA
DTNB	5,5'-dithio-bis-2-nitrobenzoic acid
dTTP	Deoxythmidine triphosphate
dUTP	Deoxyuridine triphosphate
EDTA	Ethylene Diamine Triacetic Acid
EMM	Yeast Edinburgh minimal media
GCL	Glutamate cysteine ligase
GCLC	GCL catalytic subunit
GCLM	GCL modifier subunit
GFP	Green fluorescent protein
GPx	Glutathione peroxidase
GR	Glutathione reductase
GS	Glutathione synthetase
GSH	Glutathione
GSSG	Glutathione disulphide (oxidised GSH)
GST	Glutathione S-transferase
hGS	Human glutathione synthetase
HS	Heat shock
JNK	c-Jun N-terminal protein kinase
kb	Kilobase
LB	Lysogeny broth
M	Molar
ml	millilitre
mg	milligram
MOPS	3-(n-Morpholino)Propanesulfonic Acid
mtDNA	Mitochondrial DNA
NADP	Nicotinamide adenine dinucleotide phosphate

NADPH	Reduced NADP
NFκB	Nuclear factor κB
ng	Nanogram
Nrf 1/2	Nuclear factor related factor 1 or 2
OD	Optical density
OH	Hydroxyl
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase PCR
S2	*Drosophila* Schneider 2 cells
SDS	Sodium Dodecyl Sulphate
SEM	Standard error of the mean
SFM	Serum free medium
SOC	Super optimised catabolite media
SOD	Superoxide dismutase
SSA	Sulphosalycylic acid
SSC	Sodium citrate buffer
TAE	Tris acetate buffer
Taq	Taq polymerase enzyme
TBE	Tris borate buffer
TCA	Trichloroacetic acid

TE	Tris-EDTA buffer
TNB	5-thio-2-nitrobenzoic acid
TNF-α	Tumor necrosis factor α
Trx	Thioredoxin
TrxR	Thioredoxin reductase
U	Unit
µg	microgram
µl	microlitre
UAS	Upstream activating sequence
X-Gal	5-Bromo-4-chloro-3-indolyl β-D-galactoside
YE	Yeast extract media

1 Chapter 1: Introduction

Aerobic organisms produce toxic reactive oxygen species (ROS) during normal metabolism as a result of the incomplete reduction of O_2. The free radical theory of ageing proposes that the cumulative damage caused by endogenous and exogenous ROS results in cellular senescence and ultimately death, and is supported by a well documented correlation between oxidative stress resistance and longevity. Cellular antioxidant defences have therefore been studied as a means to increase the oxidative stress resistance, and thereby the lifespan of organisms. The tripeptide γ-glutamylcysteinylglycine, commonly known as glutathione, is one of the most abundant intracellular thiols and is part of an integrated antioxidant system that protects cells from oxidative damage. Glutathione (GSH) is synthesised in a two-step reaction, with the first step catalysed by glutamate-cysteine ligase (GCL) and the second step catalysed by glutathione synthetase (GS). This chapter presents an overview of the reactive oxygen species which cause oxidative stress, and the cellular response to those stresses, with a particular focus on the antioxidant glutathione. In addition, the review provides an overview of the oxidative stress in ageing which links the damage caused by ROS to the rate at which organisms age. Finally, the use of *Drosophila* as a model organism will be discussed in relation to the manipulation of expression of antioxidant genes with a view to extending lifespan.

1.1 Oxidative stress

Oxidative stress is the situation where there is a serious imbalance between the rate of ROS production and the antioxidant defence system. This may arise from a reduced level of antioxidants, or from an increased production of ROS (Halliwell and Gutteridge, 1999; Abdollahi et al., 2004; Girardot et al., 2004). Free radicals are molecules or atoms containing unpaired electrons, therefore seeking stability by attacking molecules nearby to obtain another

electron, and resulting in damage to the structure and function of the oxidised molecule (Abdollahi et al., 2004). Antioxidants are substances which delay or inhibit oxidative damage to a target molecule by donating electrons to free radicals (Abdollahi et al., 2004).

Exposure to low levels of oxidative stress can lead to an adaptive response in a cell, where transcription of genes involved in response to oxidative stress is upregulated (Halliwell and Gutteridge, 1999). Excessive oxidative stress can lead to cell injury, and eventually death through apoptosis or necrosis (Girardot et al., 2004). High levels of ROS production are associated with cancer, artherosclerosis, neurodegerative diseases (Bains and Shaw, 1997; Droge, 2002), and they also play a role in ageing (Harman, 1956; Phillips et al., 1989; Cho et al., 2003), as listed in Table 1.1.

Table 1.1: Diseases in which oxidative stress is involved in the pathophysiology (Phillips et al., 1989; Abdollahi et al., 2004; Dugan and Quick, 2005; Reliene and Schiestl, 2005)

Disease Category	Diseases
Autoimmune	Rheumatoid arthritis, inflammatory bowel diseases
Eye	Cataract, age-related macular degeneration, retinopathy, cystic macular degeneration
GI tract	Hepatitis, pancreatitis, colitis
Kidney	Renal failure, renal interstitial fibrosis, neuropathy
Lung	Bronchial asthma, cystic fibrosis, pneumonia, pulmonary fibrosis, Chronic obstructive pulmonary disease
Neurodegenerative	Parkinson's, Huntington's, amyotrophic lateral sclerosis, Alzheimer's, multiple sclerosis, dementia, neuronal lipofuscinosis
Red blood cells	Sickle cell disease, anaemia, neonatal hypoxia, thalassaemia, malaria infection
Skin	Contact dermatitis, atopic dermatitis, psoriasis, vitiligo
Vascular	Atherosclerosis, myocardial infarction, stroke
Systemic	Cancer

1.1.1 The chemistry of free radicals

Atoms usually exist in the state where electrons are matched, or paired in terms of spin, the quantum number describing the two possible orientations of the magnetic moment generated by an electron as a result of rotation on its own axis (Cowan, 1981). This means that the outer

orbital of an atom will have two electrons, one with a spin value of ½, and the other with a value of -½ (Cowan, 1981).

A free radical is any species capable of independent existence that contains one or more unpaired electrons in the outer orbital (Figure 1.1), and they are usually highly reactive because they readily give up or accept an electron to stabilise the unpaired electron (Halliwell and Gutteridge, 1999). This leads to a chain of radical reactions, which can be quenched by a molecule that is stabilised by the addition of a free radical (Halliwell and Gutteridge, 1999).

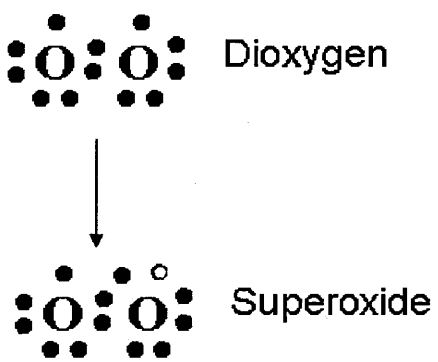


Figure 1.1: Schematic representation of the gain of electrons to create a free radical. The Lewis electron configuration of molecular dioxygen is shown, as well as the structure of free radical superoxide with the extra electron, which gives superoxide the negative charge.

The free radical status of a molecule is indicated by a dot, for example $O_2^{\bullet}$, which can exist with a positive, negative or neutral charge. The reactions in which free radicals are involved are oxidative, where there is either a gain in oxygen, e.g. $C+O_2 = CO_2$, or a loss of electrons, e.g. $Na = Na^+ + e^-$, or $O_2^{\bullet-} = O_2 + e^-$ (Halliwell and Gutteridge, 1999). Because there are examples of reactive species in biological systems which are not radicals, for example peroxide, the term 'reactive oxygen species' (ROS) is more widely used to incorporate oxygen radicals ($O_2^{\bullet-}$ and $OH^{\bullet}$) and some non-radical derivatives of oxygen, including H_2O_2, hypochlorous acid (HOCl) and O_3, which are involved in oxidative damage of the biomolecules (Halliwell and Gutteridge, 1999).

1.1.2 *Sources of reactive oxygen species in vivo*

ROS are an inevitable consequence of normal aerobic metabolism, and are produced during aerobic respiration when electrons are transferred along the mitochondrial respiratory chain to generate ATP (Sampayo *et al.*, 2003), as illustrated in Figure 1.2.

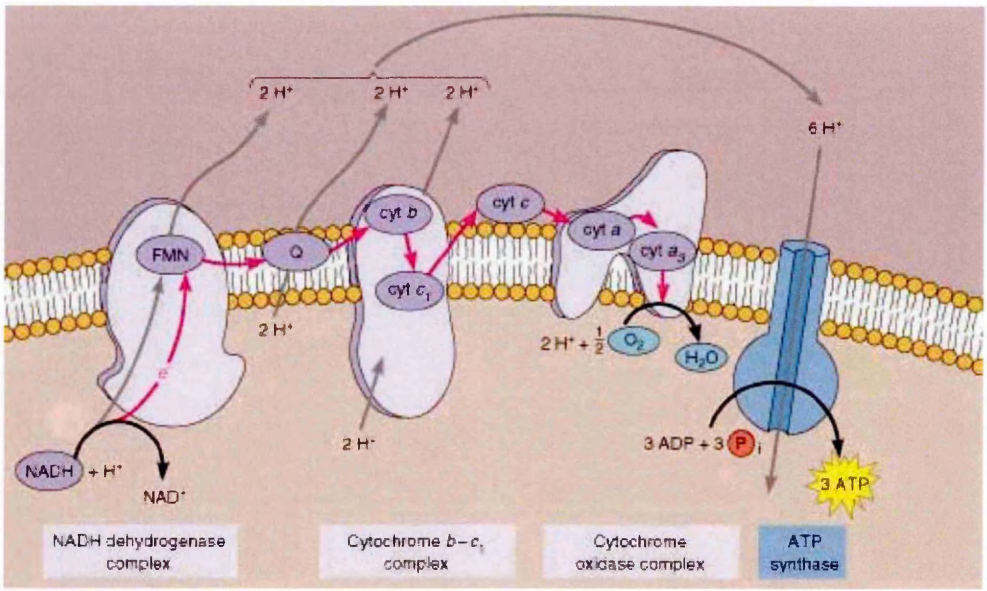


Figure 1.2: A cartoon of the components of the electron transport chain, with the movement of the electron shown with a pink arrow, and the proton pump shown with grey arrows. FMN = flavoprotein, Q=ubiquinone, cyt = cytochrome. (Becker et al., 2000)

The respiratory chain is not, however, a perfect system, and electrons are able to 'leak' out and form ROS (superoxide, hydrogen peroxide or hydroxyl molecules) when electrons are donated inappropriately, usually to ubiquinone (Phillips et al., 1989; Sampayo et al., 2003). It has been estimated that approximately 2-4% of total oxygen consumed during electron transport is converted to ROS in aerobic organisms (Nygren et al., 2005). ROS are also formed during peroxisomal fatty acid metabolism and cytochrome P450 reactions (Seto et al., 1990; Hekimi and Guarente, 2003). Cellular ROS, particularly the hydroxyl radical, are highly reactive and cause damage to lipids, proteins, carbohydrates and nucleic acids (Phillips et al., 1989; Seto et al., 1990). Biological systems also produce ROS by interaction with ionising and UV radiation, and by the redox cycling of cellular components such as quinines, thiols and flavins

(Phillips et al., 2000). The three major endogenous ROS species are discussed in the following sections.

1.1.2.1 Superoxide radical $O_2^{\bullet-}$

Superoxide is produced by the reaction: $O_2 + e^- = O_2^{\bullet-}$, and has one unpaired electron (Halliwell and Gutteridge, 1999). Sources of superoxide in aerobic organisms include the mitochondrial electron transport chain, the cytochrome P450 system in the endoplasmic reticulum and auto-oxidation reactions (Hekimi and Guarente, 2003; Nygren et al., 2005). Of these sources of superoxide, the mitochondrial electron transport chain in the inner membrane of the mitochondrion, illustrated in Figure 1.2, is the most important source of $O_2^{\bullet-}$ *in vivo*.

1.1.2.2 Peroxide ion O_2^{2-}

Peroxide is produced mainly in the mitochondrial outer membrane by the reaction: $O_2^{\bullet-} + e^- = O_2^{2-}$ (Cadenas and Davies, 2000). Peroxide is not technically a radical, as it has no unpaired electrons, but the strength of the O_2 bond decreases with added electrons, therefore the addition of two electrons to create peroxide readily leads to the formation of the radical $O_2^{\bullet-}$ (Halliwell and Gutteridge, 1999).

1.1.2.3 Hydroxyl radicals $OH^{\bullet}$

The extremely unstable hydroxyl radicals are generated by different pathways in the cell, including Fenton chemistry, where Fe^{2+} reacts with H_2O_2 to form $OH^{\bullet}$ and OH^-; UV-induced homeolytic fission of the O-O bond in H_2O_2; and ionising radiation, which can produce hydroxyl radicals from water (Phillips *et al.*, 1989). The highly reactive hydroxyl radicals are responsible for most of the oxygen-derived cytotoxicity.

1.1.3 Cellular responses to oxidative stress

The cellular consequences of exposure to highly reactive radicals are summarised in Figure 1.3. In order to combat the deleterious effects of ROS, cells have antioxidant defences available to them which vary between cell and tissue types. Within biological systems, an antioxidant is defined as a substance which delays or prevents oxidation of an oxidisable substrate, when present at levels lower than the substrate (Halliwell and Gutteridge, 1999). The antioxidant defences available to cells include ROS scavenging chemicals, enzymatic systems, and proteins such as heat shock proteins (Orr and Sohal, 2003; Sampayo and Lithgow, 2004; Nygren et al., 2005).

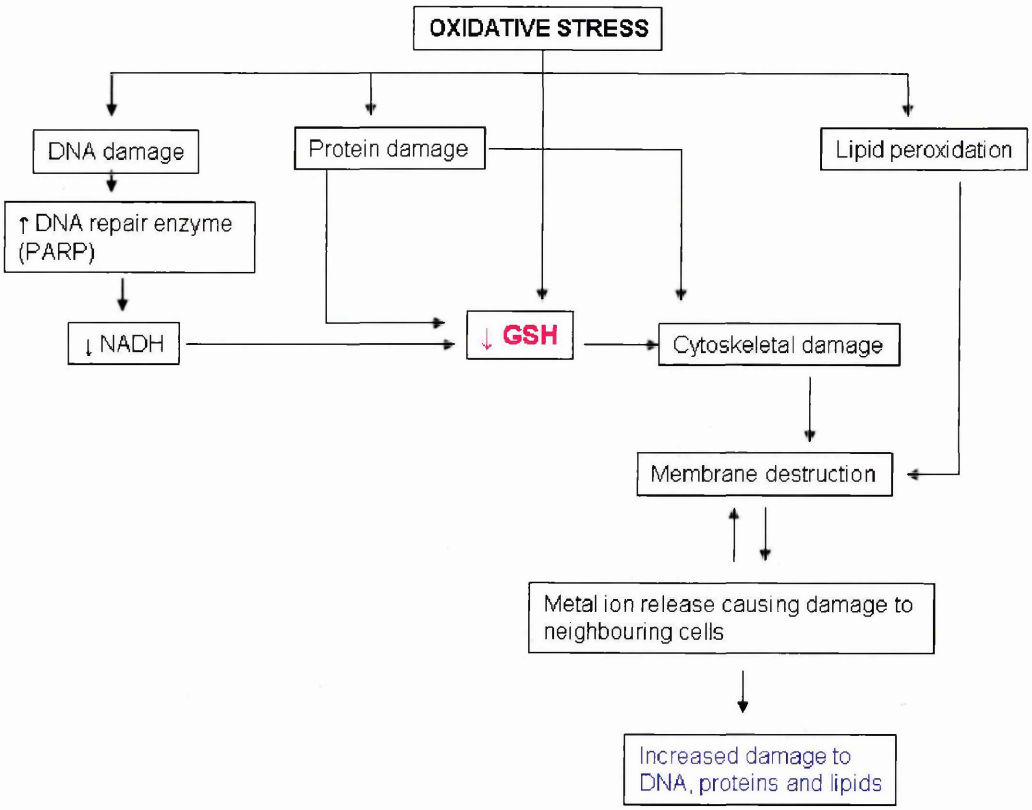


Figure 1.3: The effects of oxidative stress on cellular processes, including DNA damage, protein damage and lipid peroxidation. GSH is depleted directly and indirectly by exposure to oxidative stress. Reduction in GSH has further consequences which increase the cellular damage (Modified from Halliwell and Gutteridge, 2001).

1.1.3.1 Glutathione

The cell has a wide variety of non-enzymatic antioxidant molecules available which have the ability to remove reactive oxygen species, thus minimising damage caused by ROS to membranes, DNA and proteins. The water-soluble tripeptide glutathione (GSH) is the most abundant non-enzymatic antioxidant (Meister, 1994; Rebrin et al., 2004), and is supported by other ROS scavenging molecules including thioredoxin (Kanzok et al., 2001), selenoproteins (Morey et al., 2003) and antioxidants from the diet such as Vitamins C and E (Halliwell and Gutteridge, 1999). The functional redundancy and cooperation between the antioxidant detoxification pathways shows the importance of cellular defence against ROS. GSH is of interest in ageing studies due to its antioxidant properties, and because the levels of GSH decrease with age in various species and tissue types (Kanzok et al., 2001; Rebrin et al., 2003;

Zhu et al., 2006). Depletion of GSH has been shown to activate the apoptotic pathway, resulting in cell death (Hancock et al., 2001; Cristofanon et al., 2006). In a recent study, it was shown that human blood GSH levels correlated positively with good health in all age groups (Lang et al., 2002). Cellular redox status is maintained mainly by preserving the balance between reduced GSH and its oxidized form, GSSG (Figure 1.4). The GSH:GSSG redox couple is the most important of the redox couples because the GSH concentration is about 500-1000 fold higher than thioredoxin (section 1.1.3.2) or NADPH redox systems (Filomeni et al., 2002). Oxidation-reduction and thiol-disulfide exchange reactions during exposure to oxidative stress can result in a change in the GSH:GSSG ratio, with a shift in the ratio in favour of GSSG reflecting an oxidized cellular redox status (Aw, 2003).

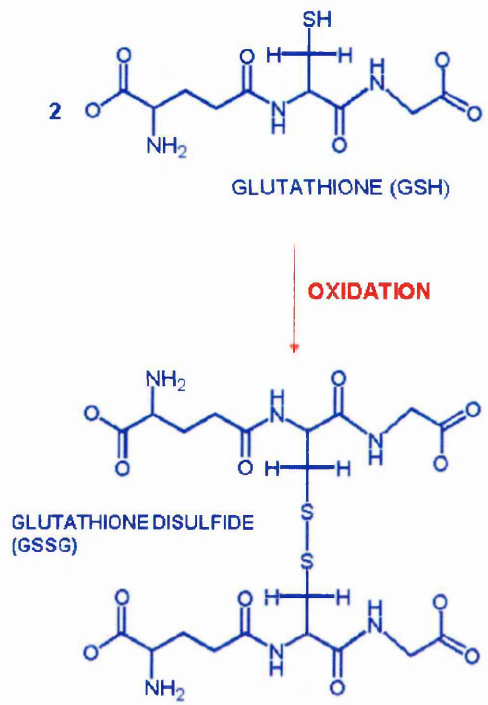


Figure 1.4: The chemical structures of reduced glutathione (GSH) and oxidised glutathione (GSSG).

Glutathione is the major non-protein thiol in many organisms, and is widely distributed in most cells. The high concentrations of GSH in most tissues, ranging from 1-10mM, as well as its high electron-donating ability make GSH a powerful cellular antioxidant (Schafer and

Buettner, 2001; Nygren et al., 2005). GSH is also one of the major determinants of cellular redox potential (Bauer et al., 2002; Cristofanon et al., 2006). The potent reducing power of GSH is the basis for its involvement in a variety of other cellular actions including detoxification, prevention of protein cross-linking, DNA synthesis, cell cycle regulation , signal transduction activity, telomerase regulation, neurotransmission, and its action as a co-factor for antioxidant enzymes (Kehrer and Lund, 1994; Meister, 1994; Halliwell and Gutteridge, 1999; Janaky et al., 1999; Polekhina et al., 1999; Meierjohann et al., 2002; Kim et al., 2003a; Borras et al., 2004; Forman et al., 2004).

The reducing properties of GSH are due to the sulfhydryl (-SH) group, also known as a thiol group, on the cysteine residue. The -SH group is the sulphur analogue of the hydroxyl group (Forman et al., 2004). The thiol exchange reactions which GSH can take part in are rapid and reversible, and are able to control redox sensitive protein functions (Cho et al., 2003). As an antioxidant, reduced glutathione is an electron donor in the neutralisation of H_2O_2 to H_2O, producing oxidised glutathione (GSSG) in a reaction catalysed by glutathione peroxidase (GPx) (Inoue et al., 1998). GPx is present in most animal tissues, but is not found in insects (Radyuk et al., 2001). To compensate for the lack of this enzyme, *Drosophila* rely on catalase, thioredoxins and peroxidases to remove peroxides (Radyuk et al., 2001), with GSH providing reducing equivalents in reactions involving peroxidases. GSH also displays antioxidant properties by reacting directly with the hydroxyl radical ($OH^{\bullet}$), reducing it to H_2O. GSH reacts directly with free radicals, but also chemically restores damage to macromolecules by donating a hydrogen atom (Nygren et al., 2005).

In an oxidation reaction, GSH is oxidised to GSSG, with the differences in structure illustrated in Figure 1.4. In healthy cells, GSSG is usually less than 10 percent of the total amount of glutathione present, and the decreased ratio of reduced to oxidised glutathione can be used as a

measure of cellular oxidative stress, redox flux and ageing (Frosali et al., 2004; Njalsson and Norgren, 2005).

A further protective role of GSH is its involvement in detoxification of electrophilic xenobiotics in a conjugation reaction catalysed by the glutathione S-transferases (Halliwell and Gutteridge, 1999; Kim et al., 2003a). GSH conjugation represents an important detoxification reaction because electrophiles are potentially toxic species that can bind to critical nucleophiles, such as proteins and nucleic acids. All the enzymes involved in xenobiotic transformation have the potential to generate reactive intermediates, most of which are detoxified to some extent by conjugation to GSH (Parkinson, 2001). GSH is also involved in organic peroxide detoxification through glutathione peroxidase (Forman et al., 2004; Zhu et al., 2006).

In mammalian cells glutathione is maintained in its reduced state by glutathione reductase (GR). This enzyme is absent in *Drosophila*, where thioredoxin (Trx) is substituted for the GR system in order to support reduction of GSSG (Kanzok et al., 2001), as illustrated in Figure 1.5. Thioredoxin reductase (TRxR) catalyses the reduction of oxidised thioredoxin to $Trx(SH)_2$, which is then involved in the reduction of other proteins such as thioredoxin peroxidase, and ribonucleotide reductase, as well as the reduction of GSSG (Kanzok et al., 2001). The thioredoxin system has been found to be capable of maintaining sufficiently high levels of GSH in *Drosophila* necessary to counteract the high concentrations of ROS present in insects (Kanzok et al., 2001; Morey et al., 2003). GSH is also exported through the cell membrane to be metabolised by γ-glutamyltransferase (GGT), a membrane bound enzyme that cleaves GSH into its component amino acids cysteine, glycine, and glutamine, which are ultimately recycled for use in gluthathione synthesis in the cytosol (Rahman et al., 2005), as shown in Figure 1.5 overleaf.

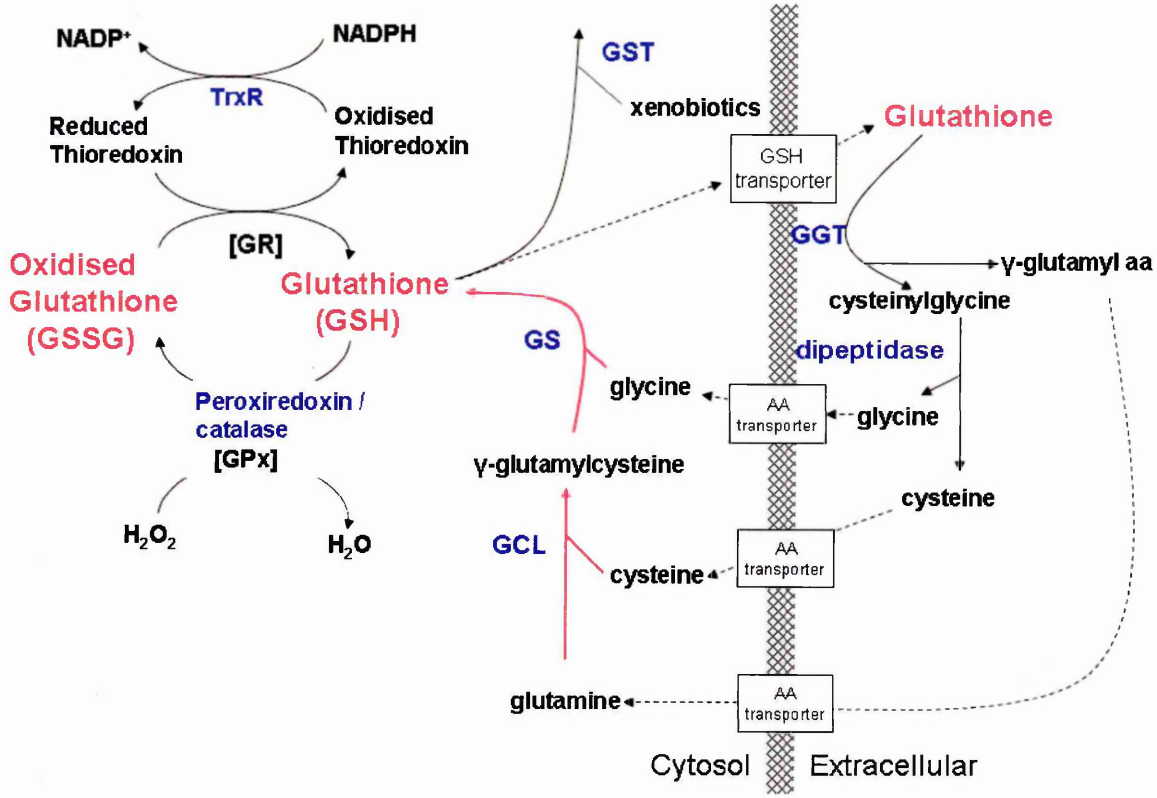


Figure 1.5: Glutathione cycle in *Drosophila*. Glutathione is synthesised in the cytosol from its component amino acids in a two-step reaction shown in red arrows, catalysed by GCL and GS. GSH is transported out of the cell by a GSH transporter, and leads to the formation of γ-glutamyl amino acids and cysteinylglycine, which is cleaved to form glycine and cysteine. Glutathione is used to detoxify xenobiotics and electrophiles present in the cell by conjugation reactions catalysed by GST (glutathione-S-transferase). Glutathione is also used inside the cell to protect the cell against reactive oxygen species, such as H_2O_2. Most eukaryotic cells utilise the GPx (glutathione peroxidase) system to detoxify peroxide, but since this enzyme is not present in *Drosophila*, flies rely on the presence of the peroxiredoxin system, as well as catalase to remove peroxides. Oxidised glutathione (GSSG) is replenished by the thioredoxin system in *Drosophila*, since flies lack the glutathione reductase (GR) enzyme. All enzymes are labelled in dark blue.

Despite the efficient renewal of GSH by GR/Trx, GSH is lost from the cell through detoxification and free radical scavenging reactions, thus necessitating *de novo* synthesis of GSH, as illustrated in Figure 1.5, and in more detail in Figure 1.6. In the first step, glutamate-cysteine ligase (GCL; EC 6.3.2.2, formerly known as γ-glutamylcysteine synthetase) catalyses the rate-limiting reaction, where cysteine is bound to the γ-carboxyl group of glutamate using ATP, resulting in the formation of and the release of ADP and P_i (Janowiak and Griffith, 2005). The formation of the gamma bond protects γ-glutamylcysteine from hydrolysis by peptidases. Availability of cysteine, produced from methionine, is usually the rate-limiting factor in this step (Meister and Larsson, 1995). GCL is a holoenzyme composed of a catalytic

(GCLC) and a modifier subunit (GCLM), which are separated under oxidising conditions (Seelig et al., 1984). *D. melanogaster* GCLC has been cloned and characterised (Saunders and McLellan, 2000; Fraser et al., 2002), and is regulated in a manner similar to mammalian GCLC, as described in 1.1.4.

In the second step of the glutathione synthesis reaction, γ-glutamylcysteine is combined with glycine by the enzyme glutathione synthetase (GS; EC 6.3.2.3) to form glutathione (Gushima et al., 1983; Meister, 1985; Huang et al., 1995). In this reaction, the transfer of the phosphate group of ATP to the carboxylate of γ-glutamylcysteine results in the formation of an unstable acylphosphate intermediate. A nucleophilic attack of this intermediate by glycine leads to the formation of glutathione, with the release of ADP and P_i (Polekhina et al., 1999; Jez and Cahoon, 2004).

Mitochondrial concentrations of GSH are similar to those in the cytosol (Cristofanon et al., 2006). GSH is transported into the mitochondria from the cytosol by dicarboxylate carrier proteins and 2-oxoglutarate carrier protein, using energy from the concentration gradient of dicarboxylic acids (Meister, 1995; Shen et al., 2005). GSH synthesis is regulated by the concentration of GSH through feedback inhibition on GCL, as well as by the intracellular level of γ-GC and the availability of its substrates; glutamate, cysteine and glycine (Meister, 1995; Meister and Larsson, 1995; Njalsson and Norgren, 2005).

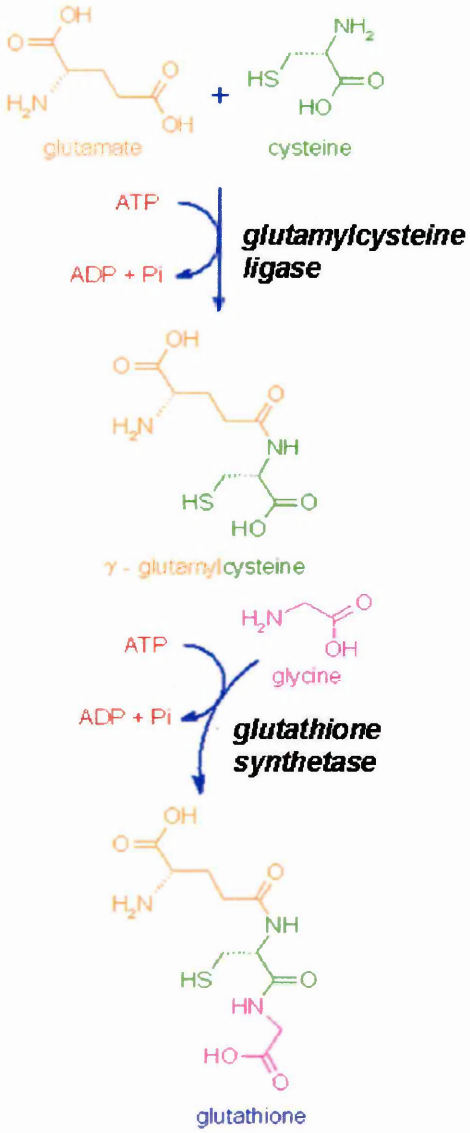


Figure 1.6: The glutathione biosynthesis pathway. The first step is catalysed by GCL, and the second step is catalysed by GS. Both steps require a molecule of ATP.

The mechanism whereby GSH regulates GCL involves the reduction of the disulphide bond between the two GCL subunits by GSH, causing a conformational change of the active site (Zhu et al., 2006). This conformational change makes the binding of GSH easier, thereby inhibiting substrate (glutamate) binding. The mechanism of glutathione biosynthesis regulation is complex (Njalsson and Norgren, 2005), and is discussed further in section 1.1.4.

1.1.3.2 Thioredoxin (Trx) and Peroxiredoxin (Prx)

Thioredoxins and peroxiredoxins are cellular reducing agents, which reduce disulphide bonds in a manner similar to glutathione (Carmel-Harel and Storz, 2000; Kanzok et al., 2001). Oxidised thioredoxin (Trx) is replenished by thioredoxin reductase (TrxR) (Missirlis et al., 2001). In *Drosophila* TrxR plays an important role in glutathione recycling, owing to the absence of glutathione reductase within *Drosophila* (Radyuk et al., 2003). Thioredoxins are also specific electron donors to peroxiredoxins (Filomeni et al., 2002).

Peroxiredoxins (Prx) are small proteins which have peroxidase activity in the presence of GSH or Trx (Radyuk et al., 2001). Both cytosolic as well as mitochondrial peroxiredoxins detoxify hydrogen peroxide, thereby protecting enzymes susceptible to oxidative damage. Five Prx members have been identified in *Drosophila*, two of which are members of the 1-Cys peroxiredoxins, which reduce H_2O_2 in the presence of glutathione, and three which have peroxide reductase activity in the presence of thioredoxin, and are known as thioredoxin peroxidases (Radyuk et al., 2001).

1.1.3.3 Antioxidant enzymes

One of the major cellular responses to oxidative stress is the use of enzymes which detoxify ROS. The main antioxidant enzymes in *Drosophila* are SOD and catalase (Mackay and Bewley, 1989; Radyuk et al., 2001; Mockett et al., 2003). Studies on transgenic animals which overexpress antioxidant enzymes, or knockout mutants which do not express a functional copy of a particular antioxidant enzyme, will be discussed later in this chapter.

Superoxide dismutase

Superoxide dismutase (SOD) is a ubiquitous enzyme that exists in two forms, encoded by two separate genes. Cu/ZnSOD, which has a copper and zinc atom in the active site of the enzyme, is present in the cytosol, while MnSOD contains manganese in the active site and is found in

the mitochondria of virtually all eukaryotic cells (Phillips et al., 1989; Orr and Sohal, 1993; Halliwell and Gutteridge, 1999). MnSOD and Cu/ZnSOD both catalyse the dismutation of $O_2^{\bullet-}$ to hydrogen peroxide and oxygen, as shown in the reaction below, thereby preventing the formation of the highly cytotoxic hydroxyl radical.

$O_2^{\bullet-} + O_2^{\bullet-} + 2H^+ \rightarrow H_2O_2 + O_2$

MnSOD and glutathione peroxidase 1 (GPx1) are the primary antioxidant defences in mitochondria (Van Remmen et al., 2004). GPx scavenges the H_2O_2 produced by SOD and converts it to water and oxygen (Van Remmen et al., 2004). *Drosophila* has been found to be well protected against the toxic effects of oxygen by SOD and catalase, which detoxifies H_2O_2, as discussed below (Seto et al., 1990). The levels of SOD are not, however, sufficient to completely eliminate superoxide molecules (Phillips et al., 2000).

Catalase

Hydrogen peroxide is formed by the dismutation of superoxide, as well as other cellular oxidation reactions, and is removed directly by catalases, while peroxidases use H_2O_2 to oxidise a reduced substrate (Mockett et al., 2003). Catalase, present in organelles called peroxisomes in most eukaryotic cells, prevents free radical reformation by breaking down H_2O_2 into water and oxygen, as illustrated in the following reaction (Halliwell and Gutteridge, 1999):

$2\ H_2O_2 \rightarrow 2\ H_2O + O_2$

There is only one functional catalase gene in *Drosophila*, and low levels of the protein are sufficient to minimise the cytotoxic effects of peroxide (Mackay and Bewley, 1989).

1.1.4 Transcriptional regulation of antioxidants

One of the mechanisms whereby oxidative stress regulates the expression of antioxidant genes is through a cis-acting enhancer sequence known as an antioxidant response element (ARE)

(Nguyen et al., 2003). The expression of a range of antioxidant genes, including glutathione biosynthesis genes (GCLC, GCLM and GS) as well as genes encoding redox proteins, e.g. thioredoxins (Yang et al., 2002; Kim et al., 2004; Rahman et al., 2005), is mediated by the presence of one or more ARE elements in the 5′ flanking region of the gene. The focus of this section is the regulation of the glutathione biosynthesis genes (Figure 1.7).

GCLC and GS are regulated in a coordinated manner, with the synthesis of GSH transcriptionally mediated by the presence of ROS (Shi et al., 1994), lipid peroxidation products (Dickinson et al., 2004), heat shock (Kondo et al., 1993), a change in the GSH:GSSG ratio (Huang et al., 2000), and by the tumor necrosis factor-α (TNF-α) signalling pathway (Yang et al., 2005b). Exposure to oxidants, such as H_2O_2, causes an initial depletion of GSH, followed by a rebound increase in GSH after approximately 24 hours, and is accompanied by an increase in GCL and/or GS transcription (Rahman et al., 2005). The regulation of GCL and GS transcription is a complex process, with a simplified schematic diagram involving some of the regulatory pathways shown in Figure 1.7.

ARE enhancer sites are present in the promoters of the genes encoding both human GCL subunits, while the gene encoding the GCLC subunit also contains an activator protein 1 (AP-1) and nuclear factor κB (NFκB) binding sites (Yang et al., 2002). The GS promoter is regulated primarily by the presence of numerous AP-1 binding sites (Yang et al., 2002), and increased AP-1 transactivation has been found to be the major underlying mechanism for the upregulation of GS and GCLC (Yang et al., 2005b).

AP-1 and NFκB are redox-sensitive families of transcription factors (Rahman et al., 2005) Both AP-1 and NFκB are activated by oxidative stress, as well as a shift in the GSH:GSSG ratio. The presence of AP-1 is essential for GS and GCL basal expression (Yang et al., 2002). AP-1 is composed of Jun and Fos proteins, and is induced by ROS via the c-Jun N-terminal protein kinase (JNK) signal transduction pathway (Rahman et al., 2005). NFκB is activated

when the depletion of GSH causes a rapid phosphorylation and degradation of the inhibitor of NFκB (IκB), illustrated in Figure 1.7 (Rahman et al., 2005).

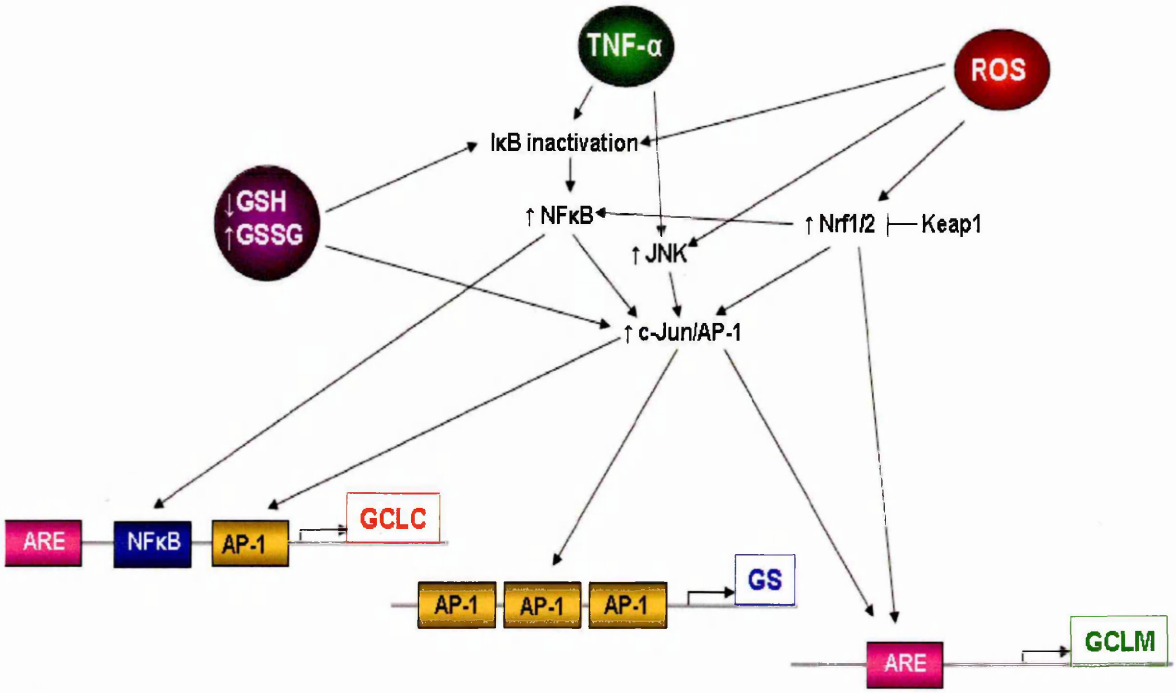


Figure 1.7: Schematic representation of the proposed mechanisms of stress-induced glutathione biosynthesis gene activation. The interactions between Nrf 1/2, NFκB, AP-1 and glutathione levels are shown in relation to ROS. The upstream enhancer elements present in the 5′ region of GCLC, GS and GCLM are shown, as well as the transcription factors acting on each of the elements.

Both ROS themselves and the end products of ROS-mediated lipid peroxidation upregulate the expression of nuclear factor related factors 1 and 2 (Nrf1 and Nrf2), transcription factors which are members of the cap'n'collar leucine zipper proteins (Iles and Liu, 2005). Nrf1 and Nrf2 are ubiquitously expressed, although expression patterns are not identical, and it is thought that Nrf2 plays a greater role in ARE-driven gene expression than Nrf1 (Mathers et al., 2004). The actin binding protein Keap1 negatively regulates Nrf by controlling its cellular localisation, as well as stability. Under reducing conditions, including DEM treatment, Keap1 is complexed with Nrf, thereby inhibiting translocation of the transcription factor into the nucleus (Itoh et al., 1999). On exposure to oxidative stress, Keap1 disassociates from Nrf1/2, allowing Nrf1/2 to move into the nucleus. Nrf1/2 transactivate ARE in the GCLC promoter,

and activate GCLM and GS expression by modulating AP-1 and NFκB expression (Yang et al., 2005a), as shown in Figure 1.7. In addition to ROS, GCL and GS are both induced by TNF-α, which acts through both AP-1 and NFκB (Yang et al., 2005b).

The transcriptional regulation of GSH biosynthesis is complex and encompasses many different signalling pathways. It was recently found that GCL is reversibly activated under dephosphorylating conditions, and in the presence of NADPH, and both processes may be used to regulate GSH production as a primary response to oxidative stress (Orr et al., 2005). The mammalian Nrf2 pathway is evolutionarily conserved amongst different genera, including zebrafish (Kobayashi et al., 2002) and nematodes (An and Blackwell, 2003). Evidence from bioinformatic studies suggests that Drosophila possess regulatory elements with functions analagous to the mammalian Nrf2 and Keap1, although the biological function of the hypothetical genes encoding Keap1 and the cap'n'collar proteins have not been experimentally determined (Mathers et al., 2004). In addition to these proposed orthologous genes, the ability of Drosophila to induce antioxidant gene expression, including GCL, on exposure to an oxidative challenge provides further evidence of the presence of an antioxidant regulatory system in Drosophila (Zou et al., 2000; Girardot et al., 2004; Mathers et al., 2004)

1.2 The Free Radical Theory of Ageing

The process of ageing, whereby organisms senesce and eventually die is a characteristic of all multicellular organisms. Ageing has been defined as 'the progressive loss of function accompanied by decreasing fertility and increasing mortality with advancing age' (Kirkwood and Austad, 2000). In *Drosophila* ageing is characterised by a decline in locomotor ability, geotaxis, learning, immune function and cardiac function (Bier and Bodmer, 2004; Johnson et al., 2006; Lesser et al., 2006; Simon et al., 2006), while mitochondria produce more ROS, and the GSH:GSSG ratio declines significantly (Sohal et al., 1995; Rebrin et al., 2004). Male

germline stem cell cycle activity also decreases with age in *Drosophila*, as do stem cells that maintain the gut (Wallenfang et al., 2006).

Ageing is an extremely complex, multifactorial process determined by genetic, epigenetic and environmental factors, and this, as well as the great diversity in lifespan among different species, has led to a variety of theories of ageing. Here the focus is on the free radical theory of ageing, which is presently the most widely accepted theory on ageing, and it underlies the work on the antioxidant glutathione undertaken in this study.

The free radical theory of ageing, first proposed by Harman in 1956, is still considered one of the most important ageing theories (Harman, 1956; Masoro, 1993; Orr and Sohal, 1993; Finkel and Holbrook, 2000; Knight, 2000; Phillips et al., 2000; Biesalski, 2002; Sohal, 2002; Harman, 2006). In the last 50 years, the free radical theory of ageing has been transformed into a more general theory which highlights the importance of ROS in ageing as a result of the gradual accrual of unrepaired oxidative damage leading to the progressive dysfunction and death of cells (Harman, 1981; Phillips et al., 2000; Sohal et al., 2002; Sampayo et al., 2003; Harman, 2006). The basic tenet of the free radical theory of ageing is that although there is a balance between ROS and antioxidant systems in healthy organisms, there will always be a small number of ROS that escape due to slightly inefficient antioxidant systems, and cause cumulative damage to biomolecules, which in turn leads to a gradual functional decline, ageing and eventually death if left unrepaired.

As discussed in section 1.1.2, normal cellular metabolism (principally mitochondrial metabolism) produces highly reactive species of oxygen, which cause damage to cellular structures unless inactivated by antioxidants. ROS cause damage to each of the four classes of macromolecules: nucleic acids, lipids, carbohydrates, and protein, and in keeping with the hypothesis of the free radical theory of ageing, damaged products of each class accumulate in ageing organisms, including *Drosophila* (Landis and Tower, 2005; Magwere et al., 2006a).

One of the most widely reported manifestations of oxidative damage to DNA is base modification, for example the formation of 8-hydroxy-2'-deoxyguanosine (8-oxo-dG) (Nygren et al., 2005). ROS also cause single strand DNA breaks, double strand DNA breaks, and interstrand crosslinks which could disrupt transcription and translation, resulting in possible cell death or cancer (Hasty et al., 2003; Kim, 2004). The estimated 10 000 oxidative DNA lesions per cell each day (Ames et al., 1993; Cardozo-Pelaez et al., 2002) highlights the scale of the potential damage to the organism.

Lipids and proteins are particularly vulnerable to oxidative damage, and membranes with a high lipid content are easily peroxidised because of the degree of unsaturation of the fatty acids making up membranes (Yu et al., 1982; Sohal and Weindruch, 1996). Biological membranes, particularly of the mitochondria and endoplasmic reticulum, are also the major site of free radical production within the cell, making the vulnerable membranes more likely targets for oxidative damage.

Proteins are particularly susceptible to ageing, owing to their structural flexibility and reactive amino acid residues (Stadtman and Levine, 2003). 'Abnormal' proteins, and protein carbonyls caused by oxidative damage accumulate with age (Sohal et al., 1995; Finkel and Holbrook, 2000; Stadtman, 2002), and age-related protein damage is exacerbated by reduced mitochondrial activity which results in lowered protein turnover rate causing increased protein half-life and damage to proteins (Landis and Tower, 2005).

1.2.1 Support for the Free Radical Theory of Ageing

In order to determine whether the damage cause by ROS to macromolecules is indeed a cause of ageing, evidence supporting the three main predictions of the free radical theory of ageing is required. The first prediction of the theory is that oxidative damage accumulates with age, the second is that treatments which lower the rate of oxidative damage result in a longer

lifespan, and the third is that interventions which increase longevity should also reduce oxidative damage (Ames et al., 1993).

Oxidative damage to proteins, lipids and DNA has been shown to accumulate with age in a variety of different tissues in *Drosophila*, worms and mice (Ames, 1989; Ames et al., 1993; Sohal et al., 1995; Sohal et al., 2002). A simultaneous increase in the rate of mitochondrial superoxide and H_2O_2 generation has also been seen in ageing insects and mammals, accompanied by a decline in antioxidant defences with age (Wei et al., 2001; Sampayo et al., 2003; Melvin and Ballard, 2006). In a recent paper, the rate of mitochondrial ROS production was correlated with lifespan potential in flies (Magwere et al., 2006a). Oxidised GSH also increases with age by up to 500% in the mitochondria of ageing rats and mice (de la Asuncion et al., 1996). In addition to these observations, microarray studies have shown an increase in heatshock proteins in ageing flies (Pletcher et al., 2002), and a downregulation of genes involved in mitochondrial metabolism (Girardot et al., 2006), suggesting an increase in the presence of damaged proteins, and mitochondrial dysfunction with age. Comparisons of ageing and oxidative stress profiles in Drosophila microarray studies reveal similar patterns of gene expression, including upregulation of heatshock proteins, immune response, and antioxidant genes (Zou et al., 2000; Seroude et al., 2002; Landis et al., 2004; McElwee et al., 2004; Park and Prolla, 2005; Girardot et al., 2006). Zou et al. compared gene expression profiles of old and oxidatively stressed flies, and found that 42 genes were co-regulated with age and oxidative stress, comprising 33% of all age-regulated genes and 18% of all oxidative stress regulated genes (Zou et al., 2000). These data strongly suggest that free radicals play a major role in regulating transcription in ageing. It must however be noted that in that particular study, 60% of age-related genes were not altered in response to oxidative stress, implying that free radicals are not the only causal factor in ageing.

The second hypothesis of the free radical theory of ageing is that lowering the rate of oxidative damage should increase longevity. This can be tested either by assessing the effect of antioxidants in the diet on lifespan, or by genetic interventions which raise antioxidant levels or reduce ROS levels. A recent study in which a meta-analysis was performed of dietary antioxidant supplement studies in humans found that although almost half of the studies showed no correlation, there was an overall significant effect of antioxidant ingestion resulting in lower DNA damage measured in white blood cells and urine (Moller and Loft, 2006). Antioxidant administration has also been shown to reverse mtDNA damage and GSH oxidation rates in mice (de la Asuncion et al., 1996; Lenton et al., 2000; Alvarado et al., 2006), as well as in *C. elegans* (Sampayo et al., 2003). Feeding *Drosophila* with antioxidant chemicals has been shown to increase lifespan in some cases (Brack et al., 1997; Bonilla et al., 2002; Kang et al., 2002), although there are criticisms of these regarding the strains used in some studies, and comparability of nutritional regimes (Sohal et al., 2002). The lifespan-extending properties of SOD mimetics have been controversial. Melov et al. reported that treatment with synthetic SOD compounds significantly extended lifespan in *C. elegans*, although subsequent studies have failed to replicate these data (Melov et al., 2000; Keaney et al., 2004) A recent study resolved some of the issues by feeding SOD mimetics to *Drosophila*, and found that the antioxidants had life-shortening effects on wild type flies and did not increase resistance to oxidative stress, although when SOD null mutants were fed the supplements, their extreme sensitivity to oxidative stress and their shortened lifespan were significantly improved (Magwere et al., 2006b). The effects of antioxidant supplementation remain controversial, and to date, no single antioxidant has produced lifespan extension in a range of long-lived strains from multiple species (Magwere et al., 2006b), although there have been many studies where antioxidant supplements do result in an extension of lifespan.

To remove the problems associated with feeding of antioxidants, including control of ingestion rate, a more widespread technique of raising antioxidant levels is the use of genetic interventions which raise endogenous antioxidant levels. Transgenic overexpression of antioxidant genes, for example SOD, has resulted in increased lifespan in mice (Harman, 2006) and flies, although the effect was only credible in flies when the genes were over-expressed in neural tissues (Parkes et al., 1998; Morrow et al., 2004), or ectopically (Orr et al., 2003), as discussed in detail in section 1.3.4.1, and Table 1.2. Overexpression of genes which reduce the level of oxidative stress have also resulted in an extension of lifespan in flies (Fridell et al., 2005), while mutations which compromise antioxidant gene function show reduced lifespans (Phillips et al., 1989).

The third prediction of the free radical theory, that there is a decrease in oxidative stress in interventions which increase longevity, is also supported by evidence from the literature. A strong correlation between increased oxidative stress resistance and longevity has been established in diverse organisms including *C. elegans*, *Drosophila*, and mammals (Lin et al., 1998; Finkel and Holbrook, 2000; Feng et al., 2001; Sohal et al., 2002; Holzenberger et al., 2003; Wang et al., 2004; Bauer et al., 2005; Berry et al., 2006). The early discovery of lifespan extension when genes such as *daf*-2 and *methuselah* are mutant (in *C. elegans* and *Drosophila* respectively) provided support for the theory because correlations between oxidative stress resistance and lifespan extension were noted in the mutants (Kenyon et al., 1993; Lin et al., 1998; Pennisi, 1998), and since then the correlation has continued to be observed.

The manner in which the free radical theory can be used to encompass or explain many existing ageing theories lends additional support. The mitochondrial theory of ageing proposes that oxidative damage to mitochondria can lead to an escalating effect where damaged mitochondria release more ROS, thereby compounding rates of oxidative damage, and eventually leading to defective mitochondria in a positive feedback mechanism (Harman,

1972; Van Remmen and Richardson, 2001). Mitochondria are particularly vulnerable to oxidative damage because of the constant production of superoxide as electrons escaping the electron transport chain combine with oxygen (Van Remmen et al., 2004). Damaged mitochondria exhibit reduced energy production, leading to compromised cell function (Van Remmen and Richardson, 2001; Sohal et al., 2002), and eventually a decline in organismal physiological function. Mitochondrial DNA is particularly susceptible to oxidative damage because ROS formed by aerobic respiration are in close proximity to mitochondrial DNA (mtDNA), and mtDNA is not protected by histones, leaving it sensitive to attack by oxidants (Cadendas and Davies, 2002). Membrane lipid peroxidation and protein degradation also leads to a loss of function, and possible increased ROS production (Van Remmen and Richardson, 2001).

The 'error catastrophe' theory which proposed that ageing is caused by translational errors that give rise to defective proteins, including proofreading enzymes involved in checking the integrity of proteins, thereby leading to an accumulation of defective proteins (Kirkwood, 1977), has recently been revisited by relating it to the mitochondrial theory of ageing (Wei et al., 2001). This revised error catastrophe theory feeds into the free radical theory of ageing which also proposes that the increased production of ROS and the result of subsequent cellular damage is a cause of ageing.

The rate of living theory is based on the inverse relationship between metabolic rate and lifespan in poikilotherms, including *Drosophila* (Pearl, 1928; Lints, 1989; Masoro, 1993). A recent study measuring the metabolic rate of *Drosophila* with age detracts from the original rate of living theory because in the study oxygen consumption did not correlate with lifespan (Hulbert et al., 2004). The initial claim of the rate of living theory that a temperature-induced increased metabolic rate would reduce the lifespan of poikilothermic animals is, however,

supported by evidence that an increased metabolic rate results in greater production of ROS, which could cause cellular damage and reduced longevity (Lints, 1989; Sohal et al., 2002). Evolutionary theories of ageing state that a reduction in the force of natural selection acting on older individuals causes ageing by allowing each generation to inherit mutations which have deleterious effects in old age (Kirkwood, 1977; Rose, 1999; Kirkwood and Austad, 2000; Rose and Long, 2002). This ties in with the free radical theory of ageing because ROS levels increase with age, as does oxidative damage of DNA, both of which are a likely cause of the mutations which accumulate in older individuals (Orr and Sohal, 1994; Sohal, 2002; Sun et al., 2002; Landis and Tower, 2005).

Caloric restriction (CR) , or more specifically, dietary restriction (DR) is the only intervention to date that causes lifespan extension in all model systems tested (Masoro, 2005), including yeast (Lin et al., 2000), worms (Hosono et al., 1989), flies (Clancy et al., 2002; Pletcher et al., 2002) and rodents (Kubo et al., 1984; Masoro, 2005). Furthermore, according to preliminary data, DR has positive effects on primate longevity (Mattison et al., 2003) . *Drosophila* is a complicated case because limiting amino acids, and not calories, extends lifespan (Mair et al., 2005; Min and Tatar, 2006). DR is thought to increase the lifespan of animals by reducing the metabolic rate (Sohal and Weindruch, 1996). DR is associated with decreased rates of mitochondrial H_2O_2 and superoxide production, and lower amounts of oxidative damage (Sohal and Weindruch, 1996; Gredilla and Barja, 2005), while antioxidants such as GSH are upregulated (Cho et al., 2003). DR also has the effect of reducing fertility, therefore it may represent a response which evolved to allow animals to survive periods of food scarcity by shutting down reproduction and diverting all energy to survival of the animal (Droge, 2005; Gredilla and Barja, 2005; Min and Tatar, 2006). The effects of caloric restriction can be interpreted in terms of the free radical theory because of the observed increased antioxidant defences, or reduced free radical production rates during dietary restriction regimes (Harman,

1981; Sohal and Weindruch, 1996; Van Remmen and Richardson, 2001; Droge, 2005; Harman, 2006; Sanz et al., 2006).

1.2.2 *Criticisms of the Free Radical Theory of Ageing*

Despite their potential for causing cellular damage, ROS are important biological molecules, and are thought to play a role in gene regulation by interacting with gene regulatory elements and transcription factors, or by mediating developmental processes controlled by incremental increases in cellular oxidation (Allen, 1993; Linnane and Eastwood, 2006). Under physiological conditions ROS are specific signalling molecules, and at low levels are essential to maintaining homeostasis (Finkel and Holbrook, 2000), acting as molecular messengers, directly regulating the transcription of certain genes, e.g. NF-κ-B, *oxyR*, and the collagen gene (Allen, 1993).

Although co-expression and altered spatial/temporal expression of SOD has resulted in lifespan extension (Orr and Sohal, 1994; Parkes et al., 1998), not all studies where antioxidants are overexpressed achieve increased longevity, as discussed further in section 1.3.4.1. An example is the simultaneous overexpression of MnSOD and catalase in the mitochondria of *Drosophila,* which significantly reduced lifespan, although the antioxidant overexpression resulted in reduced H_2O_2 production and increased oxidative stress resistance (Bayne et al., 2005). The simultaneous overexpression of both subunits of GCL in a ubiquitous pattern and at high level has also recently been shown to be lethal in *Drosophila* (Orr et al., 1995; Kotecki et al., manuscript in preparation). The deleterious effects resulting from the overexpression of certain antioxidants may be explained by the theory that a baseline level of ROS is necessary for cellular signalling, and if ROS are reduced below that level by overexpression of antioxidants the reduced rate of ageing benefits from removing ROS are outweighed by the detrimental effects of interfering with cell signalling.

Another criticism of the free radical theory of ageing is that lifespan extension by overexpression of an antioxidant gene has not been seen in mammals (Dugan and Quick, 2005). A recent paper reported on the overexpression of catalase in mice causing lifespan extension. Unfortunately the effect was reduced in later generations, and is possibly a result of founder line genes (Schriner et al., 2005). This lack of effect in mice may be because complex redox regulation of cell signalling is adversely affected in transgenic lines. It is important to note at this stage that all lifespan extension studies performed on laboratory animals are conducted in an unstressed pathogen-free environment. This lack of stress is not natural, and may minimise the effect that overexpression of antioxidant genes has on survival.

1.2.3 Summary of the Free Radical Theory

On balance of all the evidence presented in the preceding section, it seems that the free radical theory of ageing is important, but may have more relevance in organisms living under conditions of oxidative stress, as would be encountered in non-laboratory natural habitats. Under stressful conditions, the production of ROS increases, and the improvement of antioxidant systems would be beneficial for the rapid removal of the toxic species. Under normal physiological conditions or laboratory conditions where ROS production rates are relatively slow, however, the increase in antioxidants may not have any additional effect on ROS removal. It may also be detrimental to increase the levels of antioxidants far in excess of their physiological concentrations, as ROS are required in the cellular processes. Artificially decreasing the levels of ROS could even shorten lifespan by interfering with signal transduction pathways, or gene regulation, thereby masking any beneficial effects on the rate of ageing. The implications for this study are that it is likely that an increase in GS expression would probably only be noticeable under conditions of raised oxidative stress.

1.3 Genes implicated in altering lifespan

Extensive research has been carried out on the molecular mechanisms of ageing, as well as the role of genetics in ageing. The heritability of lifespan is estimated to be in the range of 10-30% (Wilson et al., 2006), with natural variation, spontaneous mutations and environmental factors also playing a role in lifespan. Microarray studies have revealed that genes implicated in lifespan are diverse, with influences on processes including insulin metabolism, developmental regulation, behavioural timing, chromosomal packaging, reproduction, signal transmission and antioxidant processes (Zou et al., 2000; Seroude et al., 2002; McCarroll et al., 2004; Melov and Hubbard, 2004; Park and Prolla, 2005; Girardot et al., 2006). The main prediction of the oxidative stress theory of ageing is that ageing cannot be slowed without the corresponding reduction of oxidative stress, and one mechanism to achieve this is to increase the cellular antioxidant concentration. The focus of this review will be on genes that have a link to altering oxidative stress and longevity, with a particular emphasis on *Drosophila.*

1.3.1 D. melanogaster as a model system for research into ageing

There have been a wide range of genes described in the literature which can affect lifespan in a variety of model organisms. Each of these organisms has specific advantages and disadvantages as experimental models. Possibly the most successful models in oxidative stress and ageing research are the invertebrates *Drosophila melanogaster* and *Caenorhabditis elegans* (Herndon et al., 2002; Kirkwood and Finch, 2002; Barger et al., 2003; Masoro, 2005), followed by rodents, and the yeast *Saccharomyces cerevisiae* (Guarente and Kenyon, 2000). *Drosophila* is a very useful model organism to use in the study of ageing, with an extensive array of molecular tools available for genetic manipulation, as well as a long research history. The 180Mb genome was sequenced in 2000 (Adams et al., 2000) and encodes approximately 13600 genes. The annotated sequence greatly facilitates in-depth genetic analyses (Bernards

and Hariharan, 2001; Adams and Sekelsky, 2002; Misra et al., 2002; Drysdale and Crosby, 2005), and has been subject to several revisions since the first release in 2000. Over 70% of genes implicated in human disease were found to have orthologues in *Drosophila*, making it a useful tool for studying the molecular and genetic basis of human diseases (Reiter et al., 2001).

Drosophila has also been useful in ageing studies because of their relatively short lifespan of ~80 days, their ease of culture, and the availability of stocks with a wide range of genetic mutations (Helfand and Rogina, 2003b). When undertaking lifespan experiments, the large numbers of genetically homogeneous populations that can be produced are useful for statistical analyses. The rate of living theory originated from experiments on the effect of temperature on longevity in flies, and most theories since have been tested in *Drosophila*. *Drosophila* is useful specifically in ageing studies as most cells in the organism, except in the gonads and parts of the gastro-intestinal tract, are post-mitotic, and therefore senescent changes are not diluted by successive cell divisions (Helfand and Rogina, 2003b; Rebrin et al., 2004).

One of the major drawbacks of using *Drosophila* in ageing research is the variation in lifespan between different strains. The genetic background of the strains used for transgenic experiments has been shown to influence the results of overexpression studies, as discussed later in this chapter. *Drosophila* laboratory strains may have assumed short lifespans by adaptation to laboratory culture conditions over many decades: this lifespan reduction may be a result of selecting for high fecundity, as well as the lack of selection to remove mutations that have deleterious effects past an age of 7 days (Sgro and Partridge, 1999; Spencer et al., 2003). Genes being tested for beneficial effects on longevity may therefore only be restoring longevity in laboratory strains to the original time period (Spencer et al., 2003).

1.3.2 Genetic techniques to study ageing in Drosophila

In order to discuss the genes implicated in lifespan extension through oxidative stress resistance in *Drosophila*, it is necessary to have an understanding of the genetic techniques employed in the experiments, and the merits of each technique. To test a proposed mechanism of ageing, transgenic animals can be used to alter the genotype of an organism and then test the resultant effect on ageing (Tower, 2000). Because ageing phenotypes are difficult to study in small organisms such as *Drosophila*, lifespan is used as a measurement of the rate of ageing. Using a decrease in lifespan is not as informative as an increase in lifespan, because genetic changes that cause decreased longevity could be doing so through a wide range of mechanisms, including new pathologies which do not normally have an influence on lifespan (Tower, 2000). Mutations that increase lifespan, however, are more likely to do so via a mechanism that is directly involved in ageing. Maximum lifespan as a measure of 'ageing' in *Drosophila* has been shown to correlate with decline in physical and mental abilities associated with ageing in a recent study where a long-lived mutant also showed decreased rates of locomotor, geotaxis and learning decline (Simon et al., 2006). For this reason, this review focuses on studies which have identified genes that cause an increase in lifespan. There are a variety of techniques available to create mutants, thus enabling research into ageing in *Drosophila.* These include more traditional genetic approaches, as well as new molecular techniques such as RNAi, as summarised in the following sections.

1.3.2.1 P-elements

P-elements are transposable elements with inverted terminal repeats, as illustrated in the general P-element structure in Figure 1.8. The elements are present in the *Drosophila melanogaster* genome and have a variety of functions in genetic manipulations of the fly, including identifying genes of interest, mutating genes, and inserting specific genes back into

the fly genome. P-elements are manipulated in M cytotype strains which lack functional P-elements. Many laboratory strains are M cytotype due to the fact that the rapid spread of P-elements in natural *Drosophila* populations occurred after most *Drosophila* strains were collected (Spencer et al., 2003).

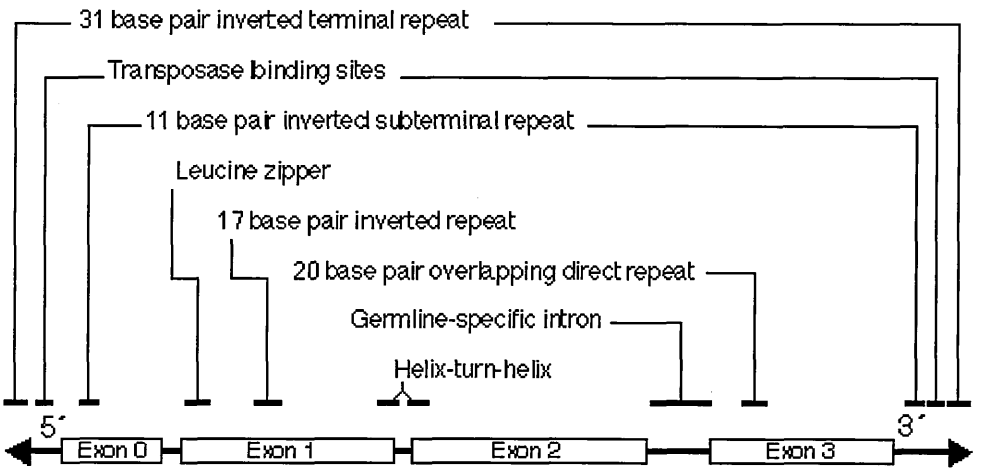


Figure 1.8: Structure of a complete P-element (Engels, 1996). The terminal repeats are indicated by arrows on either end of the structure. In the case of non-autonomous P-elements which lack their own transposase, a region between exon 1 and 3 is deleted.

Engineered P-elements that lack their own source of transposase are commonly used in genetic experiments (Ryder and Russell, 2003; Castro and Carareto, 2004) since their transposition is dependent on an external source of transposase.

When P-elements are mobilised, they often cause mutations due to insertions into genes, or as a result of imprecise excision causing deletions of the surrounding sequence. P element mobilisation occurs throughout germline development, mostly pre-meiotically, allowing the recovery of the mutations in the next generation (Engels, 1996).

1.3.2.2 Transgenes with native or heterologous promoters

In this system, the entire gene including 5′ and 3′ regulatory elements is cloned into a P-element transformation vector to produce transgenic lines. The theoretical advantage of this system is that normal regulation of expression of the gene is maintained, but the approach has

been problematic (Seto et al., 1990; Orr and Sohal, 1992; Orr W. C. and Sohal R. S., 1993; Mockett et al., 1999b), and it has recently been found that lifespan extensions are more successful when antioxidant genes are expressed using heterologous promoters (Parkes et al., 1998; Orr et al., 2003; Orr and Sohal, 2003; Orr et al., 2005). Using a heterologous promoter, transgene expression can be expressed at varying amounts and in a variety of different tissues at different lifestages. An example of a commonly used heterologous promoter is the Actin5C promoter, used for high level ubiquitous and constitutive gene expression.

1.3.2.3 The GAL4/UAS system

This system is a binary system, where the promoter activity is dependent on the presence of an activator molecule (GAL4) binding an upstream activating sequence (UAS) (Brand and Perrimon, 1993). The pUAST vector is one of the more commonly used plasmids for cloning a gene of interest in the GAL4/UAS system (Hammond, 2003), and is constructed to include a *Drosophila* promoter linked to UAS enhancer repeats. The responder strain carrying a single P-element with a transgene under UAS transcriptional control is crossed to a strain bearing a driver P-element expressing GAL4. Progeny that contain both elements will express the transgene in those cells expressing GAL4, as illustrated in Figure 1.9 (Osterwalder et al., 2001). The spatiotemporal pattern of transgene expression is dictated by the expression pattern of the GAL4 driver (Duffy, 2002). The advantage of this system is that with the wide range of GAL4 drivers available, the expression of the gene of interest can be studied in many tissue types, and toxic effects in other tissues can be avoided. It is important to note that the presence of GAL4 does not have any deleterious effects in *Drosophila* (Duffy, 2002).

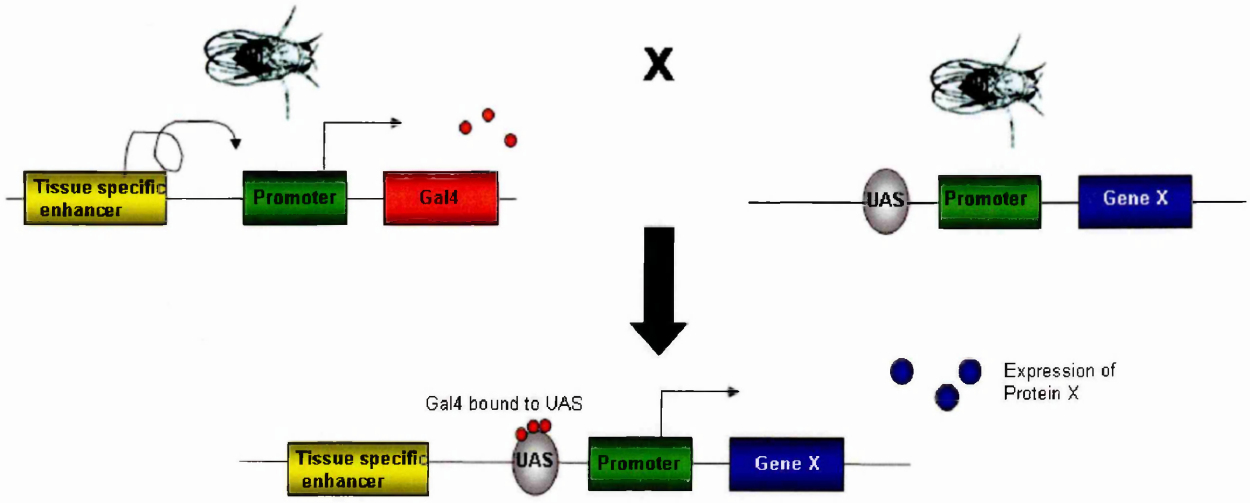


Figure 1.9: The GAL4/UAS system. The enhancer trap consisting of a P-element with a weak promoter driving Gal4 expression is present in the driver line. The UAS construct containing a promoter containing the UAS which drives expression of the gene of interest is present in the second parental stock. GAL4 is expressed and can bind to the UAS sequence in the progeny of the two lines, allowing tissue-specific expression of the gene of interest.

1.3.2.4 RU486 inducible GAL4

A tighter control of transgene expression may be achieved by using a GAL4-progesterone-receptor fusion protein (GeneSwitch) which is activated by RU486, a mammalian antiprogestin also known as mifepristone (Osterwalder et al., 2001; Roman et al., 2001). The GeneSwitch fusion protein is only able to bind to the UAS sequence of the responder line in the presence of the RU486 molecule, allowing the study of genes in adult systems which would be lethal if overexpressed during development. RU486 is easily administered to adult flies or larvae by adding the chemical to regular fly food, allowing easy control of the onset of gene expression. No adverse effects on fly lifespan or physiology have been reported for the RU486 (Roman et al., 2001; Poirier and Seroude, 2005). This system, illustrated in Figure 1.10, is favoured because the genetic background of the control flies is identical to the treated flies (which differ by the administration of RU486), and the system uses the large collection of UAS fly lines from the GAL4/UAS system, which are already characterised.

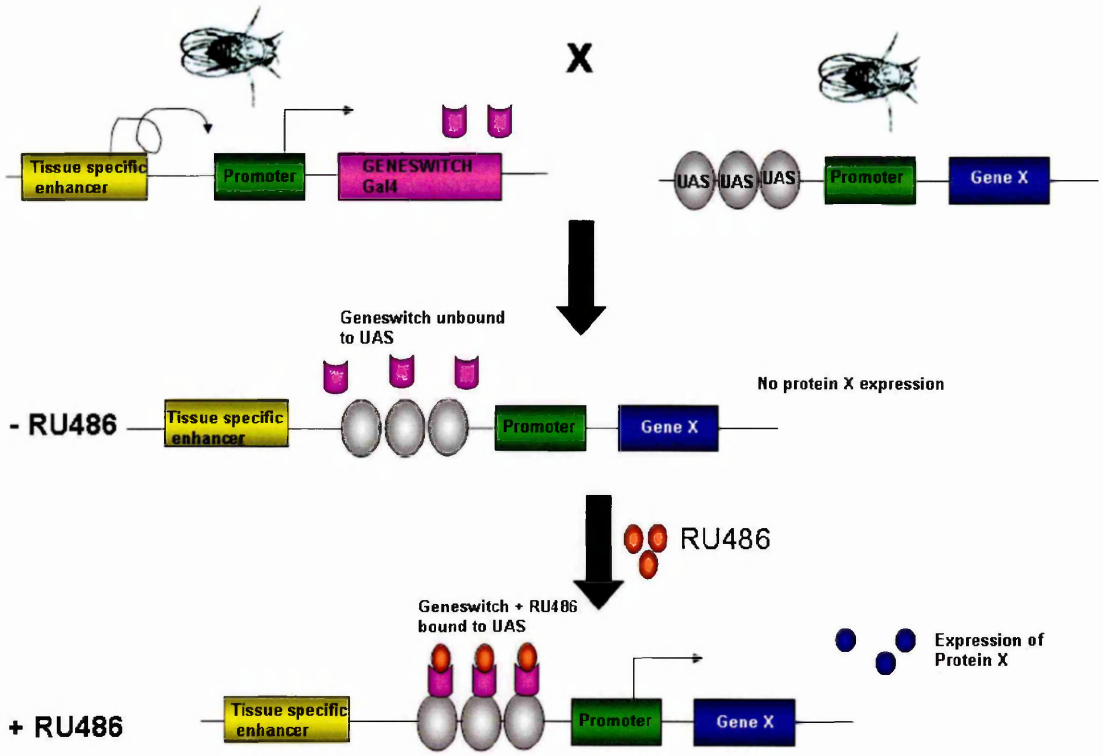


Figure 1.10: The GeneSwitch/UAS expression system in *Drosophila*. Flies expressing the tissue-specific transcriptional activator GeneSwitch are crossed to UAS reporter lines. The progeny express the GeneSwitch protein which is unable to bind to the UAS elements in the absence of RU486. In the presence of RU486, the GeneSwitch protein binds to the UAS sequence, allowing expression of the gene of interest (modified from Osterwalder et al, 2001).

1.3.2.5 FLP-out

One of the earlier inducible expression systems is the FLP/FRT binary system, which consists of one element expressing the yeast FLP recombinase under the control of the hsp70 promoter and a target element in which a stop codon upstream of the transgene ORF is flanked by FLP target sequences (FRTs) (Sun and Tower, 1999; Sun et al., 2002). Expression of FLP recombinase then excises the stop codon between the FRT sequences, resulting in the fusion of the promoter and the transgene, driving expression, as illustrated in Figure 1.11. Although this system has the advantage of removing genetic background differences from control lines, the problem with using this system in ageing studies is the requirement of the heat shock, since the heat shock itself can activate a host of stress response genes, possibly confounding the phenotype obtained from overexpression of the target gene (Sun and Tower, 1999; Sun et al., 2002).

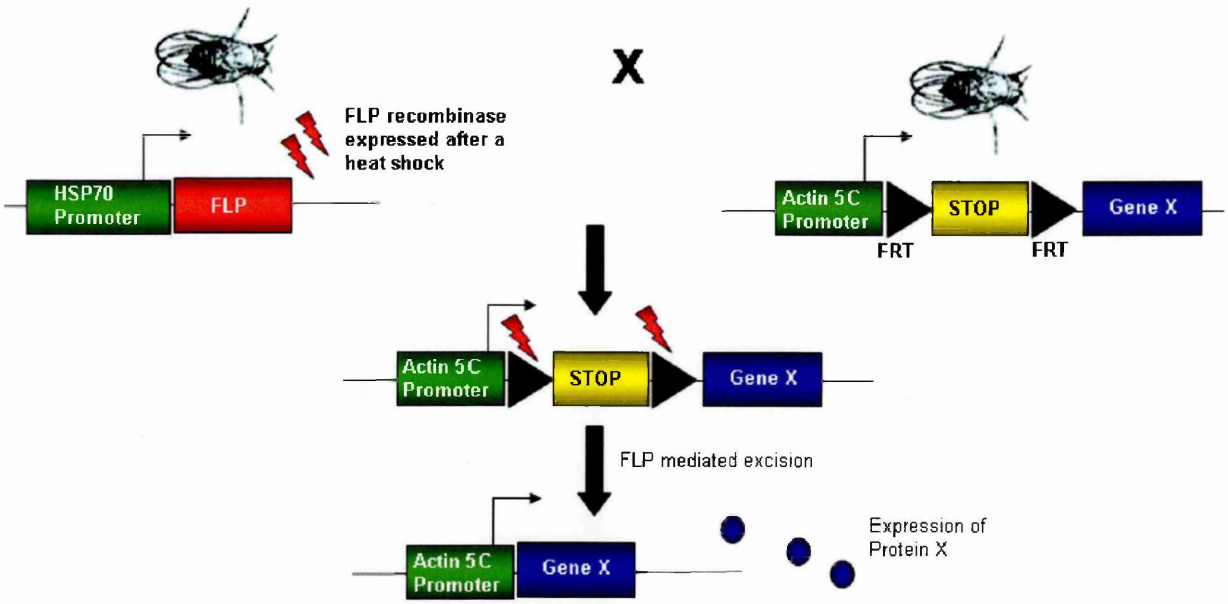


Figure 1.11 : The FLP-out system. The FLP construct consisting of yeast FLP recombinase controlled by the heat-inducible hsp70 is present in one of the parental stocks. The reporter construct is present in the second parental stock. The reporter construct contains the actin5c promoter which could constitutively drive expression of the gene of interest downstream of actin5c, but is interrupted by a transcriptional stop sequence flanked by FRTs. Upon heat shock induced expression of FLP, the FRT flanked stop sequence is excised, and the gene of interest is constitutively expressed.

1.3.2.6 RNA interference (RNAi)

RNAi is a new technique available to interfere with normal gene expression in both cell culture systems as well as in whole organisms (Clemens et al., 2000; Hammond, 2003; Lee and Carthew, 2003; Lum et al., 2003; Yang et al., 2005c). The cellular RNAi machinery can be triggered by experimentally introduced double stranded RNAs (dsRNAs). RNAi is a reverse genetic technique that degrades endogenous mRNA using specific dsRNAs. RNAi is an efficient approach to loss-of-function phenotype analyses (Boutros et al., 2004; Sledz and Williams, 2005).

RNAi is a two step process: first, an exogenous dsRNA corresponding to the target mRNA is processed by the recipient cell into 21-23bp fragments known as short interfering RNA (siRNA); this is done by the Dicer protein which functions in a similar manner to RNase III (Hannon, 2002; Amarzguioui et al., 2005). In the second step, siRNAs are incorporated into the RISC (RNA-induced silencing complex) which results in the sequence-specific

degradation of mRNA molecules from the corresponding endogenous gene, as illustrated in Figure 1.12.

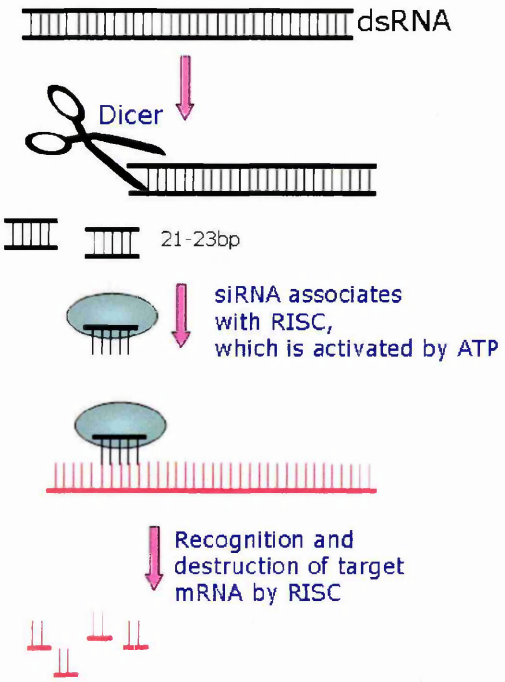


Figure 1.12: Cartoon of the mechanism of RNA interference. Synthetic double stranded RNA is shown in black, and the endogenous single stranded mRNA of interest is shown in pink.

Dicer is a member of the ribonuclease III family which contains a RNA helicase domain, a domain of unknown function, a PAZ (*Piwi/Argonaute/Zwille*) domain, two ribonuclease domains and a dsRNA binding domain (Hammond, 2005). Dicer acts on the dsRNA, cleaving it into 21-23nt small RNAs, and plays a role in delivering the siRNAs to the RISC complex (Kavi et al., 2005). *Drosophila* has two Dicers; *dcr-1*, which processes endogenously expressed microRNAs (miRNAs), and *dcr-2*, which is involved in siRNA production. The Dicer complex stabilises the siRNA before unwinding it, identifying the guide strand with the $5'PO_4$ and degrading the passenger strand that does not enter the RISC complex (Hammond, 2005; Kavi et al., 2005; Sledz and Williams, 2005).

The argonaute proteins (Ago) form part of the RISC complex, and have the characteristic PAZ and PIWI domains, which interact with Dicer and bind to the guide siRNA respectively

(Hammond, 2005). In *Drosophila* there are 5 genes encoding argonaute proteins: *piwi*, *aubergine*, *argonaute-1*, *argonaute-2*, and *argonaute-3* (Kavi et al., 2005). The RISC multi-protein components assemble the active RISC complex, into which the single stranded guide siRNA is loaded (Kavi et al., 2005). The active siRNA/RISC complex locates mRNA targets by Watson-Crick base pairing, leading to cleavage of the mRNA target by the RNase H enzyme Slicer, whose activity is contained within Ago2 (Hammond, 2005; Sledz and Williams, 2005).

1.3.3 Genetic background problems in Drosophila

It is important to control for genetic variation and its effect on lifespan very thoroughly. In addition to spontaneous mutations that accumulate in all fly lines, genetic variation is also introduced through the 'positional effect' of the random P-element insertion. Some insertion sites are close to enhancers that could have significant effects on the levels of transgene expression (Tower, 2002). P-element insertions could also disrupt a random gene unrelated to the study, causing a wide range of unforeseen effects in the animal. The use of inducible transgene constructs such as GeneSwitch, has bypassed this position effect by ensuring that control and experimental lines carry the same P-element insertion, and differ only in RU486 administration (Osterwalder et al., 2001).

It has recently been noted that there is a negative correlation between the control lifespan and the increase in lifespan of flies that overexpress a variety of antioxidant enzymes (Orr et al., 2003), as illustrated in Figure 1.13. The reason for this phenomenon is thought to be due to lack of genetic robustness in the short lived flies which is rescued by the overexpression of antioxidant enzymes. Overexpression of the same antioxidants has been shown to have a much milder, or no effect, on longevity in wild type long lived strains which are genetically robust (Spencer et al., 2003). This highlights the importance of using a long lived control fly when determining a gene's effect on longevity.

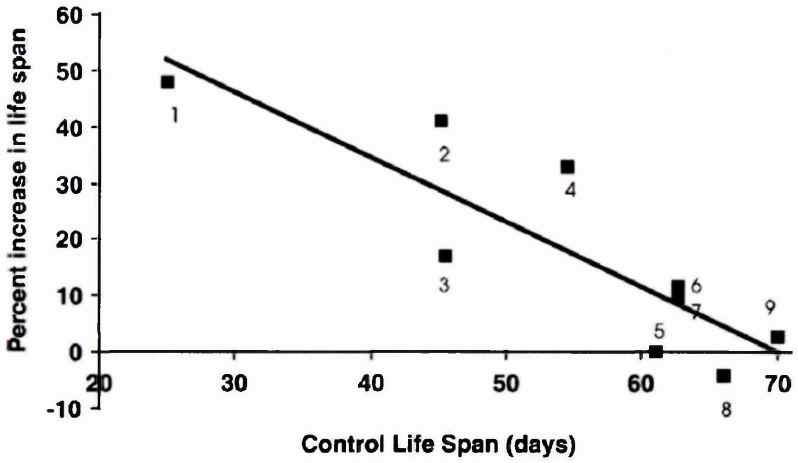


Figure 1.13: A significant correlation between lifespan extension and control lifespan in *Drosophila melanogaster* was found using data obtained from published studies featuring overexpression of antioxidant enzymes in transgenic *Drosophila* up to 2003 [(1) Sun and Tower, 1999; (2) Parkes et al.,1998; (3) Reveillaud et al., 1991; (4) Orr and Sohal 1994; (5) Seto et al., 1990; (6) Orr and Sohal, 1993; (7) Orr and Sohal, 1992; (8) Mockett et al., 1999 (MnSOD); (9) Mockett et al., 1999 (thioredoxin reductase)]. The figure was obtained from Orr and Sohal (2003).

1.3.4 Genetic interventions which increase longevity in Drosophila

Interventions which have been shown to increase *Drosophila* lifespan are antioxidant overexpression, reduction in egg laying , reduction in temperature, and dietary restriction (Maynard Smith, 1962; Rose and Charlesworth, 1981; Mair et al., 2003). This review will concentrate on genetic interventions which cause an increase in longevity, specifically overexpression of antioxidant genes, since the free radical theory can be tested by determining whether normal lifespan may be extended by increasing antioxidant levels to a point where they minimise damage caused by ROS. Mutations which result in a decrease in lifespan do not necessarily point to genes involved in ageing; the lowered fitness could be a result of decreased fitness due to the disruption of any number of biological pathways.

1.3.4.1 Overexpression of antioxidant enzymes

The four main antioxidant enzymes in *Drosophila* are Cu/ZnSOD, MnSOD, catalase and thioredoxin reductase. The free radical theory of ageing predicts that a decrease in oxidative stress would inhibit the accumulation of oxidative damage, thus increasing the lifespan of the

animal (Helfand and Rogina, 2003a; Orr and Sohal, 2003). Expectations of the free radical theory are that an addition of antioxidants to the diet would alter the balance of oxidants and antioxidants, which would lead to increased longevity. The addition of antioxidants to food is problematic as the amounts ingested cannot be quantified, and lethality at high doses may be due to starvation as the animals reject their food (Le Bourg, 2001).

Another proposed method to test the free radical theory of ageing would be to increase the activity of antioxidant enzymes and determine the effects on lifespan. The over-expression of antioxidant enzymes *in vivo*, including catalase, Cu/Zn SOD and MnSOD is discussed in detail below.

1.3.4.2 Overexpression of Cu/ZnSOD and MnSOD

Both Cu/Zn SOD, expressed in the cytosol and outer mitochondrial space, and MnSOD, expressed in the inner mitochondrial space reduce superoxide anion to hydrogen peroxide. A deficiency in SOD leads to increased oxidative stress (Orr and Sohal, 1993), and results in a reduced lifespan and oxidative stress sensitivity in SOD null adult flies (Phillips et al., 1989). The insertion of extra copies of *Drosophila* Cu/Zn SOD under native promoter control into the fly genome has no effect on lifespan (Seto et al., 1990; Orr and Sohal, 1993; Orr and Sohal, 1994). These studies used a P-element mediated approach to introduce the coding sequence of Cu/Zn SOD into the genome, and it is possible that this method could have influenced results due to the 'positional effect', although multiple insertion lines were used to control for this effect (Orr and Sohal, 2003).

Sun and Tower used an inducible FLP-out binary system to overexpress Cu/ZnSOD under the control of a constitutive promoter (Sun and Tower, 1999). This approach did not increase lifespan significantly in most of the female lines, although male lines had an increase in lifespan of up to 48%. The lifespan increase, however, was found to be dependent on genetic background, with higher levels of enzyme overexpression needed to observe an increase in

lifespan in long-lived strains compared to short-lived strains (Sun and Tower, 1999; Tower, 2002). Muscle tissue-specific overexpression of *Drosophila* Cu/Zn SOD had no effect on lifespan (Phillips et al., 2000).

The two most successful studies of Cu/Zn SOD overexpression in *Drosophila* have used transgenic expression of bovine or human SOD, and in both studies virtually isogenic strains could be compared (Reveillaud et al., 1991; Parkes et al., 1998). Overexpression of bovine Cu/ZnSOD resulted in a moderate lifespan increase of 10% (Reveillaud et al., 1991). The study using the GAL4-UAS binary system to overexpress human Cu/Zn SOD (hSOD) specifically in motorneurons showed increases in lifespan of up to 41%, although the controls were once again short-lived flies (Parkes et al., 1998; Phillips et al., 2000). Criticisms of this work that should be noted are firstly that the motorneuron 'specific' D42-GAL4 driver used also expresses GAL4 in salivary gland tissue in our laboratory (R. Saunders, pers. comm.). Secondly, the D42-GAL4 driver used has been shown to have an exceptionally long lifespan (C.Kotecki, pers. comm.), and these data were not included in the Parkes publication.

The overexpression of human SOD, using the same transgene construct as Parkes et al (1998), in wild caught long-lived flies showed that of ten lines tested, one showed extended lifespan in males, while 6 showed extended lifespan in females (Spencer et al., 2003). These results highlighted the importance of genetic background in lifespan experiments, and also pointed to an interesting sex-specific variation. On average, however, the increase in SOD expression in motorneurons increased lifespan in many genetic backgrounds, including long-lived flies (Spencer et al., 2003).

In addition to an effect on longevity, human or bovine Cu/Zn SOD overexpression resulted in increased oxidative stress resistance (Reveillaud et al., 1991; Parkes et al., 1998), although this trend was not seen when native *Drosophila* SOD was overexpressed (Seto et al., 1990; Orr and Sohal, 1993). An interesting trait noted by Orr and Sohal (1993) was that SOD

overexpression at moderate levels led to increased oxidative stress resistance; however, a severalfold increase in expression caused cells to became sensitised to paraquat induced oxidative stress, probably due to the enhanced rate of peroxide generation (Orr and Sohal, 1993). This is an important point to note when overexpressing specific antioxidants, as the raised level of one antioxidant may interfere with the delicate balance between antioxidant systems, for example, between SOD and catalase in this instance.

MnSOD over-expression under the control of a native promoter led to a 5% lower lifespan, and no increased resistance to stress (Mockett et al., 1999a), while the targeted overexpression of MnSOD in motor neurons extended lifespan by 30% (Phillips et al., 2000). The induced overexpression of MnSOD under the control of actin5C promoter in adult flies also extended lifespan by 16-33% (Sun et al., 2002).

It is possible that optimal levels of SOD are present in adult flies, and overexpressing native SOD has no effect in an already long-lived genetic background. Under conditions of oxidative stress, however, the optimal level of SOD is raised, and an increased level of antioxidant has a beneficial effect. In contrast to this, the results of overexpressing SOD using non-native promoters show that SOD does limit *Drosophila* lifespan. This may be caused by new expression patterns and significantly higher levels of the antioxidant in particularly vulnerable tissues, such as motorneurons. The positive results achieved with heterologous promoters may also be the result of detrimental effects of overexpression during development when using a native promoter (Sun et al., 2002).

On balance, it would seem that overexpression of SOD extends lifespan, however only when expressed in specific tissues.

1.3.4.3 Overexpression of catalase

Catalase mutants with a 50% reduction in catalase activity show no increased susceptibility to hydrogen peroxide stress, although mutants with 5% of wild type catalase activity are sensitive

to hydrogen peroxide stress (Mackay and Bewley, 1989). The over-expression of catalase under the control of a native promoter (Orr and Sohal, 1992), as well as a transgene under the control of an inducible actin5C promoter (Sun and Tower, 1999; Sun et al., 2002), had no effect on lifespan, although resistance to H_2O_2 induced stress was slightly increased. Overexpression of catalase targeted ectopically to the mitochondria increased resistance to oxidative stress, however, the increased levels of mitochondrial catalase did not have any effect on *Drosophila* lifespan (Mockett et al., 2003). Similarly, catalase overexpressed in motorneurons had no effect on lifespan (Phillips et al., 2000). Overexpression of catalase in *Drosophila* therefore can increase resistance to oxidative stress, although there are no beneficial effects on ageing or lifespan. The work done on catalase in *Drosophila* to date suggests that there is a threshold level of catalase activity above which viability is not affected. It was recently suggested that ectopic overexpression of catalase in mouse mitochondria extended lifespan, however the results are not conclusive because the mice used in the study were not backcrossed, and later generations did not display the lifespan extension (Dugan and Quick, 2005; Schriner et al., 2005).

1.3.4.4 Combined SOD/catalase overexpression studies

The overexpression of Cu/Zn SOD in conjunction with catalase has been found to increase *Drosophila* lifespan, although only when overexpressed in particular tissues and under the control of non-native promoters (Orr and Sohal, 1994; Sun and Tower, 1999). The first study where native SOD and catalase were co-expressed found 8 transgenic lines which had an increased lifespan (14-34%), while 1 line had decreased lifespan, and another 6 lines remained unchanged (Orr and Sohal, 1994). Genetic background in this study was a problem as the mean lifespan of control flies was 55 days, approximately 20% less than average lifespan, as illustrated by point 4 in Figure 1.13 (Orr and Sohal, 1994). In the study by Sun and Tower, the

FLP expression system used was later found to be leaky, confounding results obtained for control flies (Sohal et al., 2002).

In order to clarify the effects of antioxidant co-expression on lifespan in a long-lived genetic background, different combinations of transgenic antioxidants with normal spatial and temporal expression were simultaneously overexpressed, including Cu/Zn SOD and catalase; and Cu/Zn SOD, MnSOD, catalase (Orr et al., 2003). Overexpression of any of the combinations of these and other homologously expressed antioxidant enzymes had no effect on longevity in relatively long-lived genetic backgrounds. Further studies revealed that MnSOD overexpressed with Cu/ZnSOD had moderately additive effects (Sun et al., 2004), and confirmed that the overexpression of catalase along with Cu/ZnSOD (Orr and Sohal, 1994) or MnSOD (Sun et al., 2002) had no additional beneficial effect on lifespan extension. In fact, when catalase was simultaneously overexpressed with Cu/ZnSOD in motorneurons, the life extending effect of Cu/ZnSOD overexpression was abrogated (Parkes et al., 1998). Further, although the co-expression of MnSOD and catalase in mitochondria showed an increased resistance to oxidative stress, the simultaneous overexpression of MnSOD and catalase caused a significant reduction in lifespan via an unknown mechanism (Bayne et al., 2005).

These studies on the effect of antioxidant enzyme overexpression suggest that increasing the levels of antioxidants is only beneficial when a new antioxidant is introduced into a region of the cell or organism where it is not normally expressed, rather than increasing the activity of an enzyme that is normally present (Mockett et al., 2003). Moreover, it is likely that ROS can be removed to a baseline level with beneficial effects as a result of reduced oxidative damage, however the further reduction of ROS then has detrimental effects on lifespan due to disruption of the important cell signalling role of ROS (Bayne et al., 2005).

1.3.4.5 Other genetic interventions that increase *Drosophila* lifespan

In the previous two decades ageing research was shaped by the discovery of long-lived mutants of *C. elegans* (Friedman and Johnson, 1988; Kenyon et al., 1993; Yanase et al., 2002) and *Drosophila* (Lin et al., 1998; Rogina et al., 2000; Clancy et al., 2001). These early studies pose problems in the search for the underlying mechanism of ageing because of the multitude of trade-offs for the increased longevity, including sterility and dwarfism (Sohal et al., 2002; Spencer et al., 2003; Mockett and Sohal, 2006). The control lifespans were also short, and the lifespan extension reported for the *Drosophila* mutant *Chico*, for example, was significantly reduced in strains with a genetic background conferring increased longevity (Spencer et al., 2003). The search for a single gene that could control the rate of ageing has since been replaced by the understanding that ageing is a multi-gene phenomenon that is influenced by environmental factors. In addition to oxidative stress, a major focus in ageing research is the insulin signalling pathway. Down-regulation of this pathway has been shown to extend the lifespan in worms, flies and mice, although there is usually a reduction in fertility, and an associated dwarf phenotype (Liang et al., 2003; Tatar et al., 2003; Holzenberger et al., 2004; Baumeister et al., 2006). Another problem with pursuing this pathway is that in humans a decrease in insulin signalling is associated with diabetes, which is linked to pathologies which decrease lifespan (Rincon et al., 2004). A summary of the genes involved in the insulin signalling pathway, along with other genetic interventions which could not be discussed due to the constraints of this review, are presented in Table 1.2.

Table 1.2: Review of genetic interventions which have increased lifespan in *Drosophila melanogaster*, excluding the antioxidant gene overexpression discussed in the previous section.

Gene	Function	Intervention	Effect on lifespan	Reference
methuselah (*met*)	Transmembrane protein involved in stress response pathway.	P-element insertional disruption of gene	Up to 35% increase in lifespan. This large extension in lifespan has since been refuted, and the effect shown to be temperature sensitive (Mockett and Sohal, 2006).	(Lin et al., 1998)
chico	Insulin receptor substrate	Heterozygous or homozygous null mutations	48% mean lifespan extension in homozygous females, 13% in heterozygous males, but homozygous male lifespans were reduced. Sterility in mutants. Short-lived controls.	(Clancy et al., 2001)
Indy	Exchanger of dicarboxylic and tricarboxylic Krebs cycle intermediates.	Heterozygous null mutation	50% increase in maximum lifespan. Recently refuted, with only 15% increase in long lived background (Spencer et al., 2003)	(Rogina et al., 2000)
Drosophila Insulin Like Peptides (*Dilp*)	Signalling molecule in the insulin signalling pathway.	Ablation of neurosecretory cells where DILPs are produced	Reduction in insulin signalling results in up to 35% increase in maximum lifespan. Short-lived controls.	(Broughton et al., 2005)
InR	Insulin Receptor homologous to *C. elegans daf-2*	Hypomorphic mutants	Up to 85% increase in lifespan, but with reduced fertility and dwarf phenotype. Short-lived controls.	(Tatar et al., 2001)
dFOXO	Forkhead transcription factor, negatively regulated by the insulin signalling pathway.	Overexpression in adult fat body	20-50% increase in median lifespan, and 19% increase in maximum lifespan. Reduction in fecundity.	(Giannakou et al., 2004; Hwangbo et al., 2004)

Gene	Function	Intervention	Effect on lifespan	Reference
dSir2	Histone deacetylase, involved in gene silencing, particulary in the insulin signalling pathway.	Ubiquitous and neuronal overexpression	Up to 57% increase in median lifespan and 25% increase in maximal lifespan.	(Rogina and Helfand, 2004)
Rpd3	Histone deacetylase	Heterozygous loss of function mutants	Up to 50% increase in median lifespan.	(Rogina et al., 2002)
puckered (*puc)*	JNK-specific phosphatase that downregulates the JNK pathway.	Heterozygous loss of function mutants	Significant increases in mean and maximal lifespan.	(Wang et al., 2005)
dPOSH	Unknown function, involved in JNK signalling.	Overexpression in neurons	14% increase in mean lifespan.	(Seong et al., 2001)
hep	JNK kinase causing activation of JNK (JNK is involved in the stress signalling pathway. JNK is a kinase that phosphorylates a variety of transcription factors).	Overexpression in neurons	50% increase in mean lifespan, 25% increase in maximum lifespan.	(Wang et al., 2003)
Gclc	Catalytic subunit of GCL, the rate limiting enzyme in glutathione biosynthesis.	Overexpression in neurons	Up to 50% increase in mean and maximum lifespan.	(Orr et al., 2005)
Gclm	Modifier subunit of GCL, the rate limiting enzyme in glutathione biosynthesis.	Ubiquitous overexpression	24% increase in mean lifespan.	(Orr et al., 2005)
MsrA	Methionine sulfoxide reductase A is involved in repair of damaged proteins by reducing methionine sulfoxide to methionine.	Overexpression in neurons	Up to 70% increase in median lifespan.	(Ruan et al., 2002)

Gene	Function	Intervention	Effect on lifespan	Reference
heat shock proteins (*hsp*)	Removal of damaged proteins usually caused by heat or oxidative stress.	Overexpression of hsp22 in motorneurons	30% increase in mean lifespan.	(Morrow et al., 2004)
		Overexpression of hsp26, hsp27 and hsp68 using hsGAL4/UAS construct	30% increase in mean lifespan.	(Wang et al., 2004)
Dm*p53*	Transcription factor regulating cell cycle, with tumor suppressor function.	Dominant negative expression in adult neurons	Maximum lifespan increased by 15%.	(Lee et al., 2003)
uncoupling proteins (UCP)	Mitochondrial carrier proteins which cause uncoupling of the mitochondrial electron transport chain.	Overexpression in neurons	Up to 28% increase in median lifespan.	(Fridell et al., 2005)
Ecdysone receptor	Steroid hormone	Knockout mutation	Up to 40% increase in median lifespan.	(Simon et al., 2003)

1.3.4.6 Is antioxidant overexpression beneficial?

While the overexpression of a single, cytosolic antioxidant under the control of a native promoter has not yet been shown to increase lifespan , co-expression of antioxidants and altered spatial/temporal expression has yielded lifespan extension (Orr and Sohal, 1994; Parkes et al., 1998). Motorneurons have been implicated as a vital tissue for lifespan extension in *Drosophila* (Parkes et al., 1998; Ruan et al., 2002; Morrow et al., 2004; Bauer et al., 2005), with the implication that motorneurons are particularly sensitive to oxidative stress and ageing.

A large proportion of the antioxidant overexpression studies published to date have involved SOD, and it has been suggested that the mixed results obtained are due to optimal levels of

SOD present already in wild type flies (Mockett et al., 1999a; Magwere et al., 2006b), and raising the levels of SOD would have no further effect on lifespan. On the other hand, beneficial effects are observed when flies overexpressing antioxidants are exposed to oxidative stress (Orr and Sohal, 1994; Parkes et al., 1998). This limits the potential of raised antioxidant levels to increase lifespan, although in an environment of high stress organisms may benefit from the raised antioxidant levels. A recent example is the overexpression of hsp70 preserving muscle function in old mice without increasing lifespan (Broome et al., 2006), which implies the extension of lifespan 'quality', without an increase in 'quantity' of life. This idea of increasing lifespan quality has recently been extended to *Drosophila*, where the lag before the mortality rate increases is an indication of increased quality of life measured by maintenance of body weight, physical activity and reproduction later into the flies lives (Ruan et al., 2002).

The final point to note is that a dramatic reduction in ROS as a result of overexpressing antioxidants at a high level could have detrimental effects by interfering with cell signalling. On balance, it would appear that a moderate increase in antioxidant levels in tissues where they are limited may have a beneficial effect on the quality of lifespan, even if the actual maximum lifespan is not altered. This is interesting with respect to human health where we are constantly exposed to oxidative stress via diet and an increasingly polluted environment.

1.4 Characterisation of GS

As described in section 1.1.3.1, glutathione (GSH) is synthesised in a two step process involving two separate enzymes. The first step, which in vertebrates is the rate limiting step, is the ligation of glutamate and cysteine, catalysed by γ-glutamylcysteine synthetase (GCL) and the second step involves the addition of glycine to the dipeptide γ-glutamyl-cysteine to form GSH.

GCL is feedback inhibited by GSH and is not saturated at normal levels of cysteine, therefore increased levels of cysteine can promote GSH synthesis under certain conditions.

GSH depletion has been associated with a range of human diseases including Parkinson's disease, Alzheimer's disease, liver damage, amyotrophic lateral sclerosis, cystic fibrosis and ageing (Anderson, 1998; Schulz et al., 2000; Townsend et al., 2003). There is substantial evidence for the importance of GSH and GSH-metabolising enzymes *in vivo*. Firstly, a decrease in cellular GSH through the use of chemical GSH depletors, e.g. buthionine-S-sulfoximine (BSO) or diethyl maleate (DEM), causes an increase in the number of intracellular ROS, and severe depletion of GSH causes widespread mitochondrial damage, and results in lethal organ damage in newborn animals (Meister, 1995; Armstrong and Jones, 2002; Njalsson, 2005; Hong et al., 2006). Loss of glutathione is also thought to lead to apoptotic signalling events (Schulz et al., 2000; Shen et al., 2005), eventually causing cell death. Secondly, evidence for GSH importance is shown by defects in GSH metabolism. Humans with the rare condition of GCL deficiency, and the less rare deficiency in GS, have symptoms that include neurological defects and haemolysis, as discussed further in section 1.4.5.

GSH levels decrease with age in many organisms, including fruit flies, mice, rats and humans (Sohal and Weindruch, 1996; Schulz et al., 2000). Ageing *Drosophila* exhibit an exponential decline in the GSH:GSSG ratio, which is regulated by the temperature at which the flies are

grown, suggesting that the ratio and cellular redox state is a determinant of the physiological age of flies, and thus may be involved in determining the rate of ageing (Rebrin et al., 2004). In ageing rat brains, GSH has been shown to decrease by up to 25%, with a corresponding decrease in the GSH/GSSG ratio (Zhu et al., 2006). There is also a significant pro-oxidising shift in the glutathione redox state in ageing mice, which is thought to be a result of increased mitochondrial ROS production (Rebrin et al., 2003). The mechanisms of age-related decline of GSH are not understood, but may be due to increased demand from higher levels of ROS, or from a decrease in de novo synthesis. In rats there is no age-related decline in GR activity, while there is a decrease in GCL and GS with age, suggesting that the decrease in GSH is only due to a decrease in de novo GSH synthesis (Liu et al., 2004). A recent study of rat brains contradicts the GS findings, and reported no change in GS activity with age, although there was a drop in GCL activity (Zhu et al., 2006). This discrepancy in results deserves further study.

GS is moderately conserved between different organisms, however a small number of cysteine residues that form polar interactions with ATP, Mg^{2+} and GSH are highly conserved between all eukaryotic GS sequences (Polekhina et al., 1999; Rogina et al., 2000), and will be discussed further in Chapter 5. Because of the role of GSH in cellular metabolism, and the potential role in decreasing oxidative stress, the enzymes involved in its synthesis have been the subject of many studies, albeit with the focus on GCL. GS has received less attention in the literature because it is widely accepted that GCL catalyses the rate limiting step in GSH synthesis. A recent study has shown the importance of GS over-expression in achieving high levels of GSH production in *E. coli*, and GS is also vital to human life (Njalsson and Norgren, 2005; Liao et al., 2006).

In most eukaryotic organisms GS is a 104-112 kDa homodimer consisting of two identical 52-56 kDa subunits. Despite the lack of sequence homology and different molecular masses of

GS in different eukaryotic and prokaryotic species, they all share a common three dimensional fold and are members of the ATP grasp family (Polekhina et al., 1999). GS has been characterised in a variety of animal models, including yeast, *Plasmodium*, rodents and primates, reviewed below, but there have not been any studies on GS in *Drosophila* thus far.

1.4.1 E. coli

The *E. coli* GS enzyme is a tetramer of four identical 306-residue subunits (Gushima et al., 1983; Gushima et al., 1984). GS deficient *E. coli* are sensitive to diamide, which creates oxidative stress by oxidising GSH (Carmel-Harel and Storz, 2000). Elevated expression of GS in *E. coli* did not significantly increase GSH content (Gushima et al., 1983; Liao et al., 2006).

1.4.2 Yeast

S. cerevisiae GS exists as a dimer of 492 amino acids (Inoue et al., 1998). Yeast cells with a null mutation in GS did not have increased sensitivity to oxidative stress, however the cells did exhibit poor growth on minimal medium, suggesting that in *S. cerevisiae* GS may be required to produce GSH for the detoxification of harmful metabolic intermediates (Grant et al., 1997). Overexpression of GS had no effect on GSH concentration (Grant et al., 1997). The capacity for yeast to survive with low levels of GSH could be because γ-glutamylcysteine can act as a substitute in yeast, while this does not appear to be possible in mammalian cells (Meierjohann et al., 2002).

S. pombe GS is a tetramer of 491aa subunits (Kim et al., 2003b). Overexpression of GS in *S. pombe* is reported to lead to a 1.4-fold increase in GSH, and to protect the cells against heavy metals, as well as oxidative stresses (Kim et al., 2003b). GS deficient *S. pombe* grow slower in media without GSH supplements, and are sensitive to cadmium-induced oxidative stress (Mutoh and Hayashi, 1988; Carmel-Harel and Storz, 2000).

1.4.3 Plasmodium

Plasmodium falciparum GS, encoded by a gene with no introns, is a 655 amino acid polypeptide which is active as a homodimer (Meierjohann et al., 2002).

1.4.4 Rodents

In the rat kidney, GS is composed of two identical subunits of 474 amino acid residues (Meister, 1985). The enzyme has been shown to be upregulated by glutathione depletion using DEM or BSO (Huang et al., 2000). Six different mRNAs were identified in the mouse from a single copy GS gene that encoded a 474aa protein (Shi et al., 1996). The transcripts differed at the 5′ region, with only 3 of the transcripts producing active enzymes. The authors speculate that the apparently non-functional transcripts are involved in GSH regulation in different organs.

1.4.5 Primates

The human GS gene is ~32kb, consisting of 13 exons, 12 of which are in the coding region, and there is no evidence of alternative splicing (Whitbread et al., 1998). The GS protein is a homodimer consisting of two 52 kDa subunits (Polekhina et al., 1999).

GS deficiency causes a lack of feedback inhibition of GCL because no GSH is produced, resulting in an accumulation of γ-glutamylcysteine (γ-GC), a good substrate of γ-glutamyl cotransferase, which then overproduces 5-oxoproline (Figure 1.14), leading to a life-threatening acidosis (Gali and Board, 1997). The GS activity levels in patients with the condition are 1-30% of control values, and GSH levels are typically 5-20% of the controls (Njalsson and Norgren, 2005). GS deficiency has been described in 70 patients worldwide who have been divided into three groups based on the clinical manifestation (Njalsson, 2005). GS mutations are autosomal recessive, and all GS deficient patients have some level of GS activity, suggesting that a complete loss of GS would be lethal (Gali and Board, 1997). The

mild phenotype is caused by mutations only affecting enzyme stability, resulting in haemolytic anaemia because GSH is essential for cell membrane integrity and enucleated erythrocytes cannot produce more GS as it degrades. The moderately affected group develop metabolic acidosis, while the severely affected patients develop CNS dysfunction including mental retardation, and increased susceptibility to infection (Njalsson, 2005). There is, however, no clear link between the genotype and phenotype, and patients with the same mutation do not exhibit the same symptoms, possibly due to unique environmental and/or genetic conditions in each individual (Ristoff and Larsson, 2002).

GS activity decreased logarithmically with age in the lenses of five species of primate including Old World higher primates and prosimians (Rathbun and Holleschau, 1992)

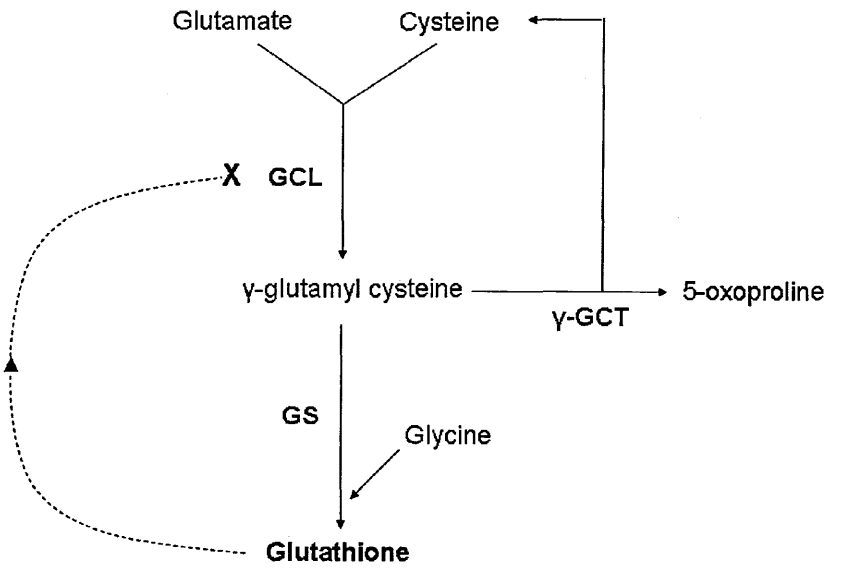


Figure 1.14: A schematic representation of 5-oxoproline synthesis within the glutathione synthesis cycle. GCL is the enzyme glutamyl-cysteine ligase, GS is glutathione synthetase, and GCT is glutamyl-cyclotransferase. In the case of a GS deficiency, γ-glutamyl cysteine (γ-GC) accumulates, which is then broken down into cysteine and 5-oxoproline by γ-GCT. (Whitbread et al., 1998)

1.4.6 Drosophila

A review of the published literature suggests that no work has been done on the characterisation of GS in *Drosophila* by other groups. The work described in this thesis is the first report describing the molecular structure of *Drosophila* GS, as well as the effects of altering GS expression in *Drosophila.*

1.5 Summary

The link between oxidative stress and longevity has been established experimentally with the extension of lifespan by mutations in genes that also decrease oxidative stress. Glutathione is a low molecular weight molecule found in most cells, and plays an important role in reducing cellular oxidative stress by detoxifying hydrogen peroxide, hydroxyl ions and xenobiotics. γ-glutamyl-cysteine-ligase (GCL) catalyses the first step of the two-step reaction to produce GSH, and the final step is catalysed by glutathione synthetase (GS). GCL is widely accepted to be the rate limiting step of the reaction, but it is thought that altering levels of GS may also impact on the amount of GSH produced, and simultaneously affect oxidative stress resistance. To date little attention has been paid to GS characterisation because it is not thought to be rate limiting. It has, however, been shown that increasing GS expression in rat liver cells raised the levels of GSH, with the authors concluding that GS may be as important as GCL in determining the rate of GSH synthesis in certain cells types (Huang et al., 2000). In *E. coli* and yeast, GS is non-essential to survival, however the absence of GS in humans is thought to be lethal. Patients with very low levels of GS have low levels of GSH, and display severe phenotypes, including mental retardation and death. A decrease in GS activity post surgery in humans was also shown to decrease GSH in skeletal muscle despite GCL activity levels remaining constant (Luo et al., 1998). Reports of GCL being the rate-limiting enzyme are concerned with hepatic systems, and it is possible that in other tissues GS may be the rate limiting step, as was the case in skeletal muscles post surgery (Yang et al., 2002). The upregulation of GS by thioacetamide in a human liver cell line also resulted in a significant increase in GSH, despite no change in GCL expression (Huang et al., 2000), while *Tat* mice exhibit decreased GSH correlated to downregulation of GS, with no change in GCL (Choi et al., 2000), again suggesting that GS plays a more prominent role than previously thought. The role of GS may therefore have been overlooked, and it possibly plays a role as important as

GCL in de novo GSH synthesis capacity, thus warranting further studies on the enzyme. The role of GS in GSH production and oxidative stress resistance in *Drosophila* is also of interest because the glutathione cycle in insects has noticeable differences compared to other species, in particular, the absence of glutathione reductase and glutathione peroxidase, thus making general assumptions of gene functions based on other species less reliable.

The broad aims of this study are to characterise the gene encoding GS in *Drosophila melanogaster*, and to determine whether resistance to oxidative stress can be manipulated by altering transcription levels of the gene encoding GS. The specific objectives include determining the structure of the gene encoding GS using molecular techniques; analysing the multiple transcripts produced by the gene through cloning and sequencing of the transcript cDNAs; determining the effect of overexpression of GS on adult fly resistance to oxidative stress; and finally elucidating the role of GS in S2 cell oxidative stress resistance and viability using RNAi.

2 Chapter 2: Methods and materials

2.1 Standard media

All concentrations of reagents described in this chapter refer to the final concentration.

Fly Media

The medium described was used for the routine maintenance of all fly stocks. 300g oatmeal, 200g molasses sugar, 48g agar and 33.6g Bakers yeast were combined in 4L of dH_2O. The mixture was boiled in a microwave oven for 15 minutes. After stirring, 13.6g p-hydroxymethyl benzoic ester in 80ml ethanol was added to the media. The mixture was then poured into plastic fly bottles or vials, and stopped with cotton wool bungs.

Fly injection plates

The following medium was used to collect embryos following egg laying by adult female flies. 50g molasses sugar, 12g agar and 8.4g Bakers yeast were combined, and made up to 1L with dH_2O. The mixture was heated in a microwave oven for 10 minutes, stirred and allowed to cool before adding 3.4g p-hydroxymethyl benzoic ester in 20ml ethanol. The mixture was then poured into Petri dishes and stored at 4°C for further use.

LB

1% tryptone, 0.5%Bacto yeast extract and 1% NaCl in dH_2O, adjusted to pH 7.2. For LB plates, agar to a final concentration of 1.5% was added to the solution before autoclaving for 20 minutes.

SOC media

2% tryptone, 0.5% yeast extract, 1% 1M NaCl, and 0.25% 1M KCl in dH_2O is autoclaved for 20 minutes. After cooling, add 1% 2M filter-sterilized glucose solution, and 1% 2M filter-sterilized Mg solution (10.15g $MgCl_2$ and12.35g $MgSO_4$ in 50ml dH_2O). Before use the entire media was filter-sterilized using a 0.22μm filter.

Yeast EMM media

EMM medium containing full supplements, except for leucine, was made up to the manufacturer's instructions: 32.5g powdered EMM medium in 1L dH_2O. The solution was autoclaved for 20 minutes. If required, 1.5% agar was added to the solution prior to autoclaving for solid media.

Yeast Extract (YE) media

0.5% yeast extract, 3% glucose and 2% agar were combined in dH_2O, and autoclaved for 15 minutes. It was necessary to ensure the glucose was completely dissolved before autoclaving to prevent caramelisation of the media.

2.2 Standard Solutions and Buffers

SSC 20X

3M NaCl and 0.3M Tri-sodium citrate in dH_2O, and autoclaved for 20 minutes before use.

Church's Buffer

0.35M Na_2HPO_4 (pH7), 0.15M NaH_2PO_4 (pH7) and 1.4% SDS are combined in dH_2O, before adding 1mM EDTA.

50X Denhardt's solution

1% Ficoll 400, 1% polyvinylpyrrolidone and1% BSA are dissolved in dH_2O over a mild heat, followed by centrifugation at 2000rpm for 5 minutes. The solution is filter sterilised with a 0.45µm syringe filter.

1M Phosphate buffer

0.18M NaH_2PO_4 and 1.55M Na_2HPO_4 in dH_2O, adjusted to pH 7.5.

10x Phosphate buffered saline (PBS)

1.37M NaCl, 26.8mM KCl, 0.1M Na_2HPO_4, 16.9mM KH_2PO_4 dissolved in dH_2O, adjusted to pH 7.4.

50X DTT

20% dithiothreitol (DTT) dissolved in dH_2O.

1M Na-EDTA

372.25g Na-EDTA dissolved in 1L dH_2O.

50X TAE buffer

2M Trizma base and 0.05M Na-EDTA in dH_2O, adjusted to pH8.0 with glacial acetic acid.

TE buffer

0.01M Trizma base, adjusted to pH 8.0 with HCl, combined with 1mM Na-EDTA in dH_2O.

10X MOPS buffer

0.2M MOPS, 0.13M Sodium acetate and 1mM Na-EDTA dissolved in nuclease free dH_2O. pH is adjusted to 7.0 with NaOH prepared in nuclease free dH_2O.

X-Gal

5% X-Gal (5-Bromo-4-chloro-3-indolyl β-D-galactoside) was dissolved in N,N-dimethylformamide

Deithly pyrocarbonate (DEPC) water

0.1% DEPC was added to 1L dH_2O and stirred overnight using a magnetic stirrer, followed by autoclaving for 20 minutes.

Antibiotic solutions

Ampicillin

Ampicillin stock solution was made up in dH_2O at a concentration of 50mg/ml, to be used at a final concentration of 50μg/ml.

Chloramphenicol

Chloramphenicol stock solution was made up in ethanol at a concentration of 50mg/ml, to be used at a final concentration of 25μg/ml.

Tetracyclin

Tetracyclin stock solution was made up in dH_2O at a concentration of 5mg/ml, to be used at a final concentration of 10μg/ml.

2.3 Enzymes

Restriction enzymes

Restriction enzymes were purchased from Amersham, Roche and New England Biolabs. Reaction conditions recommended by the manufacturer were used. For genomic DNA ~5U/μg DNA was used in the reaction, and for plasmid DNA, 2U/μg DNA was used in order to complete the digest. The reactions were typically incubated at 37°C for 3 hours, unless otherwise specified. Inactivation of restriction enzymes was performed as per the manufacturer's instructions, either by heat inactivation, or by DNA precipitation if necessary.

Ligase

Ligations were carried using a 3:1 insert to vector ratio, starting with approximately 50ng vector, to which 0.2μl (0.2 units) T4 ligase was added. The reactions were incubated at 16°C overnight.

A second Quick Ligation Kit was also used, where the ligation reaction was performed in 5 minutes compared to16 hours using T4 DNA ligase. 50 ng vector was added to a 3 fold excess of insert, along with 10μl 2X buffer and 1μl Quick T4 DNA Ligase. The reaction was incubated at 25°C for 5 minutes.

Calf intestinal alkaline phosphatase (CIP)

10 units of the enzyme was used to dephosphorylate 0.5μg digested vector at 37°C for one hour. The CIP was then inactivated by gel purification or phenol extraction.

Antarctic Phosphatase

Antarctic phosphatase was used as an alternative to CIP, using the same protocol as CIP, however the reaction time was 30 minutes rather than 1 hour, and the enzyme was heat-inactivated by exposure to 65°C for 5 minutes.

2.4 Nucleic acid extraction

Extraction of genomic DNA from adult Drosophila

The method used was obtained from FlyBase (http://flybase.bio.indiana.edu/).

Approximately 30 flies were collected into an eppendorf tube and frozen at -80°C, before homogenising them in 200µl Buffer A (100mM Tris-HCl, pH7.5; 100mM EDTA; 100mM NaCl; 0.5% SDS) with a disposable pellet pestle. An additional 200µl Buffer A was added, and the mixture was incubated at 65°C for 30 minutes. 800µl of Buffer B (5M potassium acetate, 6M Lithium chloride) was added to the homogenised flies, incubated on ice for 20 minutes, then microfuged at 13000xg for 15 minutes. 1ml of the supernatant was transferred to a clean eppendorf to precipitate the DNA with 600µl isopropanol. After centrifugation, the pellet was washed in ethanol, and resuspended in 150µl TE buffer.

Extraction of total RNA from Drosophila samples and S2 cells

The method used was obtained from FlyChip (http://www.flychip.org.uk/). RNA samples were only handled with RNase-free pipette tips and eppendorf tubes, and DEPC-treated equipment. 200µl TRIzol was added to ~15 adult flies, ~100µl larvae, ~50 µl embryos or ~$1x10^6$ S2 cells, which were then homogenised for 30-60 seconds using a disposable pellet pestle. A further 800µl TRIzol was added before centrifugation at 13 000xg for 5 minutes at 4°C. 200µl chloroform was then added to the supernatant, and incubated at room temperature for 3 minutes before centrifuging at 13 000xg for 5 minutes at 4°C. The supernatant was removed, and 0.3 volumes isopropanol was added to the pellet. After 5 minutes incubation at room temperature the sample was spun at 13000xg for 5 minutes. The resulting pellet was washed in 70% ethanol, and spun for 10 minutes. The pellet was air dried, and resuspended in approximately 100µl DEPC-dH_2O.

2.4.1 Plasmid purification

Maxipreps

Large scale plasmid purifications (for purification of ~500ug plasmid DNA) were performed according to the instructions for the Qiagen Plasmid Maxi Kit. In this protocol,

E. coli clones were grown overnight in 2L flasks containing 500ml media at 37°C in a shaking incubator. The following day, cells were centrifuged, resuspended, and lysed in an alkaline lysis solution. Proteins, genomic DNA and cellular debris were precipitated, and then removed by centrifugation. The cleared lysate was then placed on an ion exchange resin to facilitate plasmid DNA binding. A medium salt wash which removed the remaining RNA, proteins and other impurities, was followed by a high salt elution of plasmid DNA. Isopropanol precipitation resulted in a concentrated and desalted plasmid DNA pellet, which typically yielded approximately 160µg/L cells.

Minipreps

The miniprep kit used was the Wizard Plus SV Minipreps DNA Purification System. The protocol followed is detailed in the manual; briefly: a 5ml overnight culture of an *E.coli* clone was pelleted by centrifugation and resuspended. A lysis solution was added to the cell solution, as well as an alkaline protease solution, resulting in lysis of the cells, and inactivation of cellular endonucleases. The reaction was neutralised, and the solution was centrifuged to remove the cellular debris. The cleared lysate was then added to the spin column, where the DNA was bound onto a membrane during centrifugation of the lysate. After an ethanol wash, the plasmid DNA was eluted with 100µl dH_2O, typically at a concentration of 100-200ng/µl.

2.4.2 DNA extraction from an agarose gel

The DNA sample of interest was run on a low percentage (0.8%) agarose gel, and cut from the gel using a sterile blade while visualised on a UV transilluminator. DNA was then extracted from the gel fragment using a SpinPrep gel DNA kit, following the manufacturer's instructions. Briefly, the gel fragment was added to the supplied gel melt solution and heated until the agarose was dissolved. The DNA was then bound to a silica membrane in the spin tube though a series of centrifugations. The bound DNA was then washed with an ethanol solution to remove protein and salts. Elution was performed with the supplied elution solution, or with dH_2O.

2.5 Nucleic acid purification

2.5.1 Ethanol precipitation

Ethanol precipitation was performed to remove salts from nucleic acid preparations. In the case of RNA, all reagents were made up using DEPC-treated dH_2O.

0.5 volumes 0.3M sodium acetate (pH5.2) was added to the DNA/RNA solution along with 2 volumes ice cold ethanol. The mixture was incubated on ice for 20 minutes, then centrifuged at maximum speed in a benchtop microfuge for15 minutes. The supernatant was discarded, and the pellet was washed with 70% ethanol, before being allowed to air dry. Once dry, the pellet was resuspended in dH_2O or TE buffer.

2.5.2 Phenol-chloroform precipitation

Phenol-chloroform precipitation was used to remove protein from DNA extractions, and to inactivate restriction enzymes that could not be heat inactivated. A mixture of phenol:chloroform:isoamyl alcohol at a ratio of 25:24:1 was used instead of the standard phenol:chloroform mixture because the addition of isoamyl alcohol ensures the deactivation of RNases.

An equal volume of phenol:chloroform:isoamyl alcohol was added to the DNA solution, and mixed to form an emulsion before centrifuging at 12 000xg for 15 seconds in a microfuge. The upper aqueous layer was transferred to a fresh tube and the process was repeated until no white protein precipitate was visible at the aqueous-organic interface. An equal volume of chloroform was then mixed with the aqueous layer, followed by centrifugation at 12 000xg for 15 seconds. The aqueous layer was transferred to a new tube, and the nucleic acid was recovered by ethanol precipitation.

2.6 Spectrophotometric nucleic acid quantification

The amount of DNA or RNA in solution was determined by a spectrophometric measurement. The OD of a volume of DNA or RNA in TE buffer was determined at 260nm, and to calculate the concentration of the nucleic acid, the OD was multipled by the

dilution factor, and by 50 (in the case of DNA) or 40 (in the case of RNA), because an OD of 1 corresponds to 50μg/ml dsDNA and 40μg/ml ssRNA.

To estimate the purity of the sample, the OD of the sample was read at 260nm as well as 280nm, because the ratio between the OD at 260nm and 280nm gave an indication of the purity of the sample. An OD_{260}/OD_{280} value of 1.8-2 indicates a pure nucleic acid preparation, and a value of 1.5 or less indicates protein or phenol contamination of the sample.

2.7 Gel electrophoresis

2.7.1 Agarose gels

The gels were made and run in 1X TAE buffer. The concentration of the agarose ranged from 0.5-1.5%, depending on the size of fragment to be separated. 0.5 μg/ml ethidium bromide was added to the gels in order to visulaise DNA under UV light. Samples were loaded into the gel in 1X loading dye (6X loading dye contained 0.25g bromophenol blue, 0.25g xylene cyanol and 15g Ficoll in 100ml dH_2O), alongside the 1Kb plus molecular weight marker, or HyperladderII molecular weight marker. Gels were generally run at 100V for 1-2 hours.

Isolation of DNA from agarose gels was performed using the SpinPrep Gel DNA kit, as described in section 2.4.2 .

2.7.2 Denaturing gel electrophoresis of RNA

Before starting, the gel tanks, trays and combs were soaked in 3% hydrogen peroxide and rinsed with DEPC-water. RNA samples were only handled with RNase-free pipette tips and eppendorf tubes. Agarose was dissolved in 1X MOPS, and cooled before adding 18% formamide. 1X MOPS was used as a running buffer to run the denatured RNA samples. The samples were denatured in a denaturing buffer (345μl formamide, 20μl formaldehyde, 35μl 10X MOPS, 3μl ethidium bromide) by heating to 65°C. The loading buffer consisted of 50% glycerol, 40mM DEPC-treated EDTA, 0.2% xylene cyanol and 0.2% bromophenol

blue. Samples were loaded alongside an RNA size ladder, and electrophoresed for approximately 1 hour at 100V.

2.8 Subcloning DNA fragments

This technique was used when the gene of interest could not be amplified by PCR, or when the TOPO cloning vector did not contain the appropriate cloning sites (Ausubel et al., 1987).

The starting DNA molecule (in the form of a genomic DNA, a BAC or a plasmid) was digested with one or two restriction enzymes, depending on the sequence of the target DNA, and on the restriction sites available in the cloning vector. Both the insert DNA and the vector the insert would be cloned into were digested with restriction enzymes. The reaction was usually performed on 5-10μg DNA containing the insert of interest, in a volume of 20μl. The digested vector and DNA insert were then run on a low percentage agarose gel, and relevant DNA bands were cut from the gel and purified.

After gel purification, the 5′ phosphates on the digested vector DNA were removed by CIP, or alkaline phosphatase digestion, followed by an ethanol precipitation of the digested insert DNA and the dephosphorylated vector. The vector and insert DNA were then ligated using T4 DNA ligase at a vector:insert ratio of 3:1. The ligated vector was then transformed into electrocompetent *E. coli* cells, followed by colony screening to detect positive clones.

2.9 TOPO cloning

When cloning PCR products with TOPO vector, the TOPO TA cloning Kit version K2 was used. Top10 chemically competent cells with blue/white screening were transformed with linearised plasmid vector. The TOPO vector system has thymidine overhangs with topoisomerase bound to the vector. Since Taq polymerase adds an adenosine to the 3′ ends of PCR products, they can ligate to the thymidine overhanging residue in the vector. The topoisomerase binds linear ends of the vector, protecting them from digestion and prevents

religation. Ligation of the topoisomerase-activated vector to PCR products, which lack 5′ phosphates, occurs with the release of topoisomerase enzyme. Ligation of the 200ng PCR product and 2.5ng plasmid was carried out at room temperature for 30 minutes (for >5kb fragments), then added to 50µl Top10 cells. The reaction was incubated on ice for 15 minutes, followed by a 30 second 42°C heat shock. 250µl SOC was added to the cells, which were incubated at 37°C for an hour before plating them out on LB-Amp plates coated with 40µl 40mg/ml X-gal. 10 white colonies (indicating disruption of the lacZ gene, and therefore a positive cloning result) were picked and tested for the presence of the DNA insert of interest.

2.10 Bacterial transformation

Transformation was performed on electrocompetent XL1 Blue *E. coli* cells, or on One Shot Top10 electrocompetent *E. coli* cells. Transformation was carried out using a Stratagene apparatus set at 1800V. ~100ng of plasmid DNA was added to 50µl electrocompetent cells, mixed, and then added to a chilled cuvette. The cells were pulsed at 1800V before the immediate addition of 100µl SOC medium. After transferring the cells to an eppendorf tube, they were incubated at 37°C for 1 hour before plating out onto LB plates containing an antibiotic appropriate to the plasmid. In the case of plasmids carrying the lacZ gene which allows for blue/white colony screening, X-Gal was spread on the plates before addition of cells. Both XL Blue and Top10 cells are sensitive to ampicillin.

2.10.1 Creation of electrocompetent cells

Electrocompetent cells were prepared by first inoculating a colony of XL1 Blue *E.coli* into 50ml LB, and incubating at 37°C for 5-8 hours. This starter culture was used to inoculate 500ml LB which was incubated at 37°C with shaking to an OD600 of 0.5-0.7. The cells were then chilled in an ice-water bath for 15 minutes, before centrifuging at 4200rpm for 20 minutes at 2°C in a Beckman U2-21 centrifuge. The pellet was resuspended in 5ml ice-cold water, then mixed with 500ml ice cold water and centrifuged as before. The pellet

was resuspended, and washed again, and resuspended finally in a 10% glycerol solution before storing the cells in 100µl aliquots at -80°C.

The transformation efficiency of each batch of electrocompetent cells created was tested by transforming a 100µl aliquot of cells with 1ng pUC18 DNA. Volumes of 1µl, 10µl and 50µl of the transformation were plated onto LB+Amp plates, and incubated overnight at 37°C. The number of colonies on each plate was counted the following day, and the transformation efficiency was calculated as the number of colonies per µg plasmid. This calculation involved multiplying the number of colonies counted by the dilution factor, and by 1000 to convert the value from ng of DNA transformed to µg. An efficiency of over 1 x 10^8 transformants per µg plasmid in a 100µl reaction was usually obtained, but if the number was lower than 1 x 10^7, the cells were discarded.

2.10.2 Colony screening

Clones were screened in a two-stage process, first the colonies were picked and screened via PCR, after which positive clones were confirmed by restriction digest

PCR colony screening

This initial screening technique was used to detect possible positive colonies. Approximately 50 colonies, or 10 white colonies in the presence of blue/white screening, were picked using an individual sterile pipette tip for each colony. The pipette tip from each colony was dipped in 5µl dH_2O in a 10µl eppendorf tube, and then streaked on an LB plate to be incubated at 37°C overnight and stored at 4°C for further analysis. The eppendorf tube containing the *E. coli* cell solution was heated to 95°C for 10 minutes to lyse the cells. The solution was then centrifuged at 8000xg in an Eppendorf microfuge for 10 minutes to pellet the cell debris, and 3µl of the supernatant was used in a PCR reaction to test for the presence of the insert of interest. In the case of a positive PCR result, the colony was screened by restriction digest to confirm the result using the colony streak made on the LB plate.

Restriction enzyme digest colony screening

Clones with a positive result from PCR were picked using a sterile pipette tip which was placed in a 15ml sterile centrifuge tube containing 5ml LB plus the appropriate antibiotic. The culture was incubated in a shaking incubator overnight at 37°C. The following day, plasmid DNA was extracted from the culture using the miniprep procedure. Plasmid DNA was then digested using a restriction enzyme that had a recognition site within the insert DNA, resulting in a banding pattern specific to the positive clones when the digests were run on an agarose gel.

2.11 Yeast transformation

For a single transformation, 10ml EMM in a 50ml sterile tube was inoculated with a single *S. pombe* colony. The cells were grown in a shaking incubator at 25°C. The following day the OD of the solution was monitored, and when the media reached an OD of 0.4, the cells were centrifuged at 1500g for 5 minutes, and the pellet was washed in 1ml sterile dH_2O. The cells were then washed in 200µl of a 0.1M lithium acetate solution (adjusted to pH 4.9 with acetic acid). Following centrifugation, the pellet was resuspended in 50µl 0.1M lithium acetate. 600ng of plasmid DNA was added to the cells, along with 300µl of a lithium acetate/PEG solution (0.1M lithium acetate, 50% w/v PEG 4000). The cells were incubated at 30°C for 30 minutes, and heat shocked at 42°C for 15 minutes. The cells were centrifuged at 1500g for 5 minutes, and the pellet was resuspended in 200µl TE buffer. The total transformation was plated onto EMM –Leu plates which selected for the pNMT1 vector which carried the Leu2 gene.

2.12 Bacterial and yeast culture storage

E. coli bacterial cells were stored on LB plates containing the appropriate antibiotic, and *S. pombe* yeast cells were maintained on YE agar plates. The plates were sealed with Parafilm and could be stored at 4°C for the duration of a month. For long-term storage, glycerol stocks were made by mixing 850µl of a culture grown to an OD of ~0.3 with

150µl sterile glycerol. The mixture was then frozen in a dry ice bath (ethanol poured onto dry ice), and stored in a -80°C freezer indefinitely.

2.13 Vectors

The pBluescript SK^+ vector was used to clone restriction enzyme digested fragments. The TopoTA Vector 2.1 was used to clone PCR products. Cloned sequences were sequenced from this vector using the universal M13 primers.

The pNMT TopoTA expression vector was used to clone and express PCR amplified DmGS cDNAs in *S. pombe.* The vector contains a *nmt1* promoter which results in thiamine-regulated gene expression in *S. pombe* cells. It has been found, however, that the *nmt1* promoter does allow gene expression in the presence of thiamine in some cases, and the ability to 'shut off' completely seems to be gene dependent (Forsburg, 1993). For this reason the thiamine control was not used in this experiment.

The pUAST vector was used to clone restriction enzyme digested fragments for expression within the *Drosophila* adults using the GAL4/UAS system (Brand and Perrimon, 1993).

All the vectors used in this thesis are summarised below in Table 2.1.

Table 2.1: Vectors used for cloning DNA fragments. The uses and properties of each vector are listed.

Vector	Use	Size (kb)	Antibiotic resistance	Promoter	Source/ Reference
pBluescript SK^+	Cloning restriction enzyme digested genomic fragments	3.4	Chlor-amphenicol	lac promoter	Stratagene
TopoTA 2.1	TA overhangs allows cloning of PCR products. Sequencing of cloned products using M13 primers	3.9	Kanamycin Ampicillin	lacZ promoter	Invitrogen

Vector	Use	Size (kb)	Antibiotic resistance	Promoter	Source/ Reference
pNMT TopoTA	Clone and express PCR amplified cDNAs in *S. pombe.*	6.1	Ampicillin	*nmt1* promoter which results in thiamine-regulated gene expression in *S. pombe* cells. It has been found, however, that the *nmt1* promoter does allow gene expression in the presence of thiamine in some cases, and the ability to 'shut off' completely seems to be gene dependent (Forsburg, 1993). For this reason the thiamine control was not used.	Invitrogen
pUAST	Clone and express restriction enzyme digested fragments in *Drosophila*. The construct is flanked with P-element sequences to allow for transgene genome integration. Expression of the *white* gene allows for transgenic fly identification.	9.1	Ampicillin	UAS promoter which is activated in the presence of GAL4	Brand and Perrimon (1993)

2.14 Nucleic acid hybridisation assays

2.14.1 Radiolabelling DNA probes

DNA probes of PCR or enzyme digest origin were labelled by α ^{32}P dCTP incorporation, using the Megaprime DNA labelling system. This system uses random sequence hexanucleotides to prime DNA synthesis along the DNA segment. Klenow DNA polymerase 1 uses the primer-template complex as a substrate for DNA synthesis, incorporating α ^{32}P dCTP instead of an unlabelled nucleotide. The labelled probes are

released from the template by denaturation, and added directly to the hybridisation reaction.

The Megaprime reaction is started by denaturing a 50µl solution of 50ng DNA and 5µl primers at 95°C for 5 minutes. This is followed by the addition of 10µl labelling buffer, 5µl labelled dCTP and 2ul Klenow polymerase enzyme. The reaction is incubated at 37°C for 30 minutes before being denatured by heating to 95°C for 5 minutes, followed by a cooling on ice.

2.14.2 Southern blots

DNA was separated by electrophoresis on a 1% agarose gel at a low voltage overnight. The gel was then depurinated by soaking in 0.125M HCl for 10 minutes, followed by submerging in 0.4M NaOH for approximately 30 minutes. A capillary blotting apparatus was set up by placing a paper wick on a platform supported above a tank of 0.4M NaOH. The treated gel was placed on the wick, covered by a Hybond N^+ nylon membrane then three sheets of 0.4M NaOH saturated 3MM paper. The gel and membrane were finally covered by a 20cm stack of absorbent paper towel and a weight on top of the paper stack. The blotting apparatus is illustrated in Figure 2.1. After the overnight transfer, the membrane was washed in 2X SSC buffer, and stored for the hybridisation procedure.

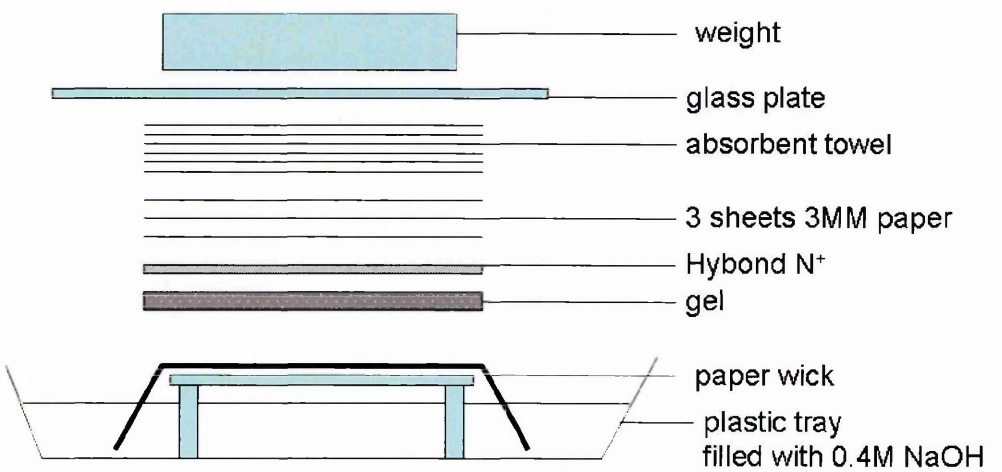


Figure 2.1: Diagrammatic representation of a capillary blotting apparatus

Probe hybridisation

Hybridisation was carried out in a Hybaid Micro4 rotating hybridisation incubator. The nylon membrane containing bound DNA was pre-hybridised with Church's buffer for 30-60 minutes. The labelled probes were then added to the Church's buffer and hybridised to the membrane in the rotating incubator overnight at 65°C. The following day the membrane was washed with 3X SSC containing 0.5% SDS at 65°C, followed by two washes with 1X SSC containing 0.5% SDS at 65°C. The membranes were then exposed to X-ray film at -80°C, and developed 1-5 days later.

2.14.3 Northern blots

All procedures for Northern blotting were done with care to avoid nuclease contamination, including rinsing apparatus in 3% H_2O_2, and making up solutions with DEPC water. RNA samples were run on a 1% denaturing agarose gel at a low voltage overnight. The gel was set up in the same manner as the Southern blot (Figure 2.1), although instead of using 0.4M NaOH, a 20X SSC solution was used to moisten the paper wicks, and to fill the lower tank. The second difference between the Northern and Southern blot set up was the use of an uncharged Hybond N nylon membrane for transfer of the RNA. After the overnight transfer, RNA was fixed to the nylon membrane using UV light in a Hoefer UVC500 UV Crosslinker.

Probe hybridisation

Northern hybridisation was carried out in a Hybaid Micro4 rotating hybridisation incubator. The nylon membrane containing bound RNA was washed in prehybridisation solution (50% deionised formamide, 5X Denhardt's solution, 1% SDS, 6X SSC) for 2 hours at 42°C. The labelled probes were then added to the pre-hybridisation solution and hybridised to the membrane in the rotating incubator overnight at 42°C. The following day the membrane was washed twice with 3X SSC containing 0.5% SDS, followed by a wash with 1X SSC containing 0.5% SDS. Both washes were performed at 55°C. The membranes were then exposed to X-ray film at -80°C, and developed 1-5 days later.

2.15 Polymerase chain reaction

The Opticon2 thermocycler was used for routine PCR as well as for RT-PCR. Thin-walled low profile PCR tubes were used in the machine. Reactions were usually carried out in 20µl volumes, containing 1X PCR buffer with Mg^{2+}, 0.2mM dNTP, 100-200ng target DNA, 0.5µM of each primer, and 1U Taq. The following PCR thermal cycling program was followed, and adjusted depending on the annealing temperature of the primers:

95°C	3 minutes	
95°C	30 seconds	
Annealing temp	45 seconds	30 cycles
72°C	1-3 minutes	
72°C	5 minutes	

Primers were designed to have an annealing temperature of 55-60°C, and care was taken to ensure that the sequences would not cause hairpin structures through complementary annealing. Primers were ordered from Invitrogen through the online ordering service at www.invitrogen.com, and the lyophilized primers were made up to a stock concentration of 100µM with dH_2O. The various primers used are listed in Table 2.2.

Table 2.2: List of PCR primers. GS distal copy on BAC=24700-33850, GS proximal copy = 34840-41280

Primer name	Sequence from 5′ to 3′	Position on BACR29H04
GS-L1	tcg ttg aga gcg cag cag ta	24970 / 35130
GS-L2	cag ccg ttt aac ggt gtc ca	24870 / 35030
GS-L3	gta tat cgc ccg gct ggt at	24740 / 34910
GS-R1	ctg ctg ctg gta tct ggg at	25190 / 35350
GS-R2	ggc ata atc ctt ggc ctt gg	25310 / 35470
GS-R3	ata cga gcc ctc act agc ga	25410 / 35570
GS-L4	cca ggc gga gat caa aag ca	23780 / 33940
GS-R4	atg gat tcg cct ctg tgc tg	26320 / 36470
GS-R5	gcg atg aga tga gat gag tg	27280 / 37440
GS-R6	tgt tat cgc aca cgc ctc ga	28330 / 38480
GS-L5	agc gag ctc tgc aac caa tc	22810 / 32440
GS-L6	gag gga cgt cga tat gtg ac	21874
GS-R7	gct cat cct cgt cat cgc ta	33670 / 41090
GS-R8	cca gac tga ggc aac agg tg	33740 / 41160
Exon1-F	ggt gga aca ggt gga gca c	24910 / 35060
Intron1-F	ggc gaa caa cta aca ac	25130 / 35280
5′ of 297-F	cag ctt gta gtc ata tcg g	28660 / 38820
GS-upstream-F	gga cca cta gtt cac gac	22940
GS-upstream-R	gaa cat gcg aac tgt gcc	23413
GS-copy2_R	gct tta gcc cca att cca g	41284
GS_copy2_notI-F	*gac taa agc ggc cgc* aca aca acg att gga cct	24520 / 34680

Primer name	Sequence from 5′ to 3′	Position on BACR29H04
GS_copy2_not1_R	*gaa tca agc ggc cgc* atc cag ttt gca tat tga	33850 / 41280
Ex4-F	cgc aag gag ttc gaa aa	25740 / 35890
Exon4-R	cgt cgt cgt gat gaa ctc c	25810 / 35970
Ex5-F	gcg atc tga tgc tgg agt c	26850 / 37005
Ex5-R	gct tcc acg agc agt aag	27087 / 37242
Ex6-F	ttt ggc cac ttg gga cca g	27134 / 37289
Ex7-F	gcc tag agt ttt cat ttg ctg ac	27695 / 37850
Ex7-R	ggt gtc ggt tgg cag gtg c	27798 / 37952
Ex8-F	gca ggt cga gat caa cac g	31830 / 39250
Ex10-F	tca ttc tgt tca tca tcg ag	32693 / 40114
Ex12-F	gac caa ggg cta tat ggt g	33450 / 40870
Ex12-R	cta ctt ctg ctc atc ctc gt	33680 / 41100
First-exon_F	cga tca tgt cca gcg acg c	25223 / 35378
Plac1	cac cca agg ctc tgc tcc cac aat	234 on EP-element
Pry4	caa tca tat cgc tgt ctc act ca	7856 on EP-element
Actin5C-F	tta ttc cag tca ttc ctt tca	
Actin5C-R	ttc gag atc cac atc tgc tg	
NMT *pombe* forward	ttt caa tct cat tct cac ttt ctg a	pNMT TOPO vector
URA4 *pombe* reverse	aca agg cat cga ctt ttt caa ta	pNMT TOPO vector
pUAST F seq	agc gca gct gaa caa gct aaa c	sequencing
pUAST R seq	ttt gtc caa tta tgt cac acc aca	sequencing
M13 F sequencing	gta aaa cga cgg cca g	sequencing
M13 R sequencing	cag gaa aca gct atg ac	sequencing
GFP-forward	gcg taa tac gac tca cta tag gga gaa tgg ctagc	GFP for RNAi
GFP-reverse	gcg taa tac gac tca cta tag gga gat tga agt tca	GFP for RNAi
GS-RNAi-F	tct gat tca tcc gcc gct gac caa g	GS for RNAi
GS-RNAi-R	cgc tat cga tca gat atg gac tgt c	GS for RNAi
GCLC-RNAi-F	ggg tct act gag cga ggg	GCLC for RNAi
GCLC-RNAi-R	ccg tgc ggc aat cca cc	GCLC for RNAi

2.16 rt-PCR

RT-PCR was performed on RNA extracted from *Drosophila* embryos, larvae, adult flies, and S2 cell cultures. In the case of embryos and larvae, approximately 100µl embryos were collected for RNA extraction, and in the case of adult flies, 10-15 flies were collected for each RNA extraction. For S2 cell culture, approximately 10^6 cells were pelleted for RNA extraction. Gene-specific primers were used to quantify the GS or GCLC mRNA in each sample after RT-PCR, using actin5C as an internal control for quantifying RNA.

2.16.1 First strand cDNA synthesis

cDNA was created from RNA using SuperScript II Reverse Transcriptase kit from Invitrogen. 2μg RNA was added to 500ng oligo(dT)$_{12-18}$ primer and 1μl 10mM dNTP solution in a total volume of 12μl. This mixture was heated to 65°C for 5 minutes then chilled on ice. After a quick spin, 5μl 5X First strand buffer, 2μl 0.1M DTT, and 200U SuperScript II RT was added to make up a final volume of 20μl. The reaction was incubated at 42°C for 50 minutes, then inactivated by heating to 70°C for 15 minutes.

2.16.2 PCR reaction

PCR from the cDNA template was optimised carefully so that the final PCR product made was still in the linear phase of the reaction. The number of cycles of the PCR reaction was decreased to the point where the reaction just became visible when visualised on an agarose gel. The concentration of primers used was tested on a gradient system, ensuring they were not in excess, which allowed for greater control of the linear phase of the reaction. The standard amount of cDNA used in each 20μl PCR reaction was 1μl. The primer concentrations and cycle numbers for each gene specific reaction are listed below:

Table 2.3: Primer concentrations and cycle numbers of specific RT-PCRs

Primer set	Primer concentration	Number of cycles
GS First ex F + Ex12R	0.4μM (fly cDNA) or 0.5μM (S2 cell DNA)	24
GCLC_F + GCLC_R	0.4μM	24
Actin5C_F _ Actin5C_R	0.4μM	21

2.17 The typhoon 9410 variable mode imager & Imagequant software

The Typhoon imager was used to scan ethidium bromide stained DNA on agarose gels, as well as peroxidase labelled western blot membranes. The ImageQuant software was used to digitally measure the density of bands on the gel or membrane, thereby allowing comparison of band intensities.

2.17.1 Measuring the intensity of DNA bands on an agarose gel

The Typhoon scanner software was used to select the size of the gel. The fluorescence acquisition mode was selected, and the pixel size was set to 200μm. After the scan was

complete, the saved image file was opened with ImageQuant software, which was used to analyse DNA band intensity on the gel. Manual analysis was performed on the gel instead of the automatic analysis to allow greater sensitivity when choosing dimensions of the bands. In manual mode, the lanes were defined, and the bands to be analysed were demarcated. The software finally subtracted the background intensity, giving values (in arbitrary units) for the volume of each band indicated on the gel. This allowed for comparison of the intensity of different bands on the same gel, eg, a comparison of GS and actin5C intensities was used to determine relative expression levels in different samples.

2.17.2 Capturing the chemiluminescent signal from a peroxidase treated western blot

The Typhoon scanner software was used to select the size of the membrane. The chemiluminescent mode was selected, with medium sensitivity set. Pixel size was set to 500μm, and the image was scanned. No further analysis of the image was performed. The scanner was used as an alternative to X-ray film to visualise the chemiluminescent reaction.

2.18 Sequencing

Samples for direct nucleotide sequencing were sent to Macrogen (www.macrogen.com). 15μl of purified plasmid (100ng/μl) or purified plasmid (50ng/μl) along with 10μl of the relevant primer (5mM) was sent to the sequencing service in Korea. The service utilised the BigDye Terminator sequencing kit from ABI, and the reactions were run on an ABI 3730XL automatic sequencer. The results were sent back in the form of .abi files. These files were then viewed and analysed using BioEdit sequence alignment editor, version 7.0.4.1 (Hall, 1999).

2.19 Glutathione determination

The principle of this assay is based on an enzymatic recycling reaction where glutathione reductase is used to quantify the amount of GSH in a sample (Tietze, 1969). 5,5′-dithio-

bis-2-nitrobenzoic acid (DTNB) is reduced by GSH in the presence of $NADPH_2$ to the yellow coloured 5-thio-2-nitrobenzoic acid (TNB) which is measured at 405nm. The mixed disulfide, GSTNB (between GSH and TNB) that is produced in this reaction, as well as GSSG, is reduced by glutathione reductase to recycle the GSH and produce more TNB. The rate of TNB production is proportional to this GSH recycling reaction, which is then directly proportional to the concentration of GSH in the sample. The reaction is started, then the colour change is monitored over 4 minutes to determine the total amount of GSH present in the samples based on the standard curve produced. The full protocol used is detailed below:

Samples were prepared by homogenizing flies or S2 cells in ice cold 10% sulphosalycylic acid (SSA), followed by centrifugation at 3000g for 10 minutes. The aqueous layer was collected and used for the assay. A standard curve was prepared using GSH in SSA ranging in concentration from 1 – 100μM.

50μl of each of the samples and standards were aliquoted in triplicate into wells of a 96-well plate. 50μl of 4mM DTNB (prepared in a 200mM phosphate buffer pH 7.5, 10mM EDTA) was added to each of the samples, followed by an addition of 0.1U glutathione reductase in phosphate buffer to each of the wells. The plates were then incubated at room temperature for 10 minutes. 50μl of 1.36mM NADPH was added to each of the samples to start the reaction, and the first reading at 405nm on a spectrophotometric multiplate reader was taken immediately. After 4 minutes, the second reading was taken.

To calculate the amount of GSH present in each of the samples, the mean rate of the reaction was calculated using the equation: Mean V = (OD at 4 min – OD at 0 min)/4

The standard curve was plotted with mean rate against GSH concentration. The total GSH concentration can then be calculated from the standard curve with the rate of colour change (mean V) proportional to the GSH standard concentrations.

2.20 Protein techniques

Protein Extraction

Extraction from yeast cells

Six millilitres of a starter culture of *S. pombe* cells was grown to an OD of 0.5. The cells were spun at 1500xg for 5 minutes, and the pellet was washed in 5ml dH_2O. The resulting pellet was resuspended in 300µl 1.85M NaOH with 7% β-mercaptoethanol. After vortexing for 1 minute, 150µl of 100% cold trichloroacetic acid (TCA) was added. The mixture was then incubated on ice, followed by a 10 minute centrifugation at 4°C. The pellet was washed in 500µl 1M Tris base, and resuspended in 150µl of an 8M urea solution. The protein sample was then stored at -20°C until required.

Extraction from adult flies

15 flies were homogenised in 150µl lysis buffer (50mM Tris-HCl, pH7.2; 20mM Na_2HPO_3, pH7; 150mM NaCl; 1% Triton X-100, 0.1mM benzamidine; 0.1mM PMSF, 0.5mM EDTA), spun in a microfuge at top speed for 3 minutes. The supernatant containing the protein was collected and stored at -20°C for further use.

Protein Determination

Protein was quantified using the BioRad DC protein assay. This assay is based on the Bradford method where Coomassie brilliant blue dye is added to a protein sample and the change in absorbance is read at 595nm. The absorbance maximum of the dye changes from 465nm to 595nm on binding to protein. Comparison of samples to the standard curve gives a relative measurement of protein concentration. The kit was used according to the manufacturer's instructions, as detailed below.

BSA protein standards were made from 0-1.4 mg/ml. The standards were treated in the same manner as the samples. 5µl of samples and standards were pipetted into individual wells of a 96-well microtitre plate. 20µl Reagent S was added to 1ml Reagent A, and 25µl of this mixture was added to each well, followed by 200µl of Reagent B. After a fifteen minute incubation, the samples were placed onto a plate reader to measure the absorbance

at 595nm. Using the BSA standard curve results, the protein concentration in each sample could be calculated.

SDS-PAGE

Protein samples were reduced prior to loading on a SDS-PAGE gel by the addition of a final concentration of 1X dithiothreitol (DTT), and boiled at 95°C for 15 minutes with 1% bromophenol blue. Approximately 20μg protein was separated on a 10% SDS-PAGE gel with a 5% stacking gel, alongside a rainbow molecular weight marker. The mini-PROTEAN II gel pouring and electrophoresis apparatus was used for the protein gels. Samples were electrophoresed in running buffer (25mM Tris, 192mM glycine, 0.1% SDS) at 100V for approximately 1 hour.

The SDS-PAGE gel consisted of two portions; the lower 'running gel' (50% Protogel acrylamide [30% acrylamide and 0.8% bis-acrylamide]; 0.15% SDS; 0.56M Tris-HCl, pH 8.8; 0.25% TEMED; 0.15% APS) which made up the majority of the gel, and a 'stacking gel' (1.67% Protogel acrylamide; 0.1% SDS; 0.125M Tris-HCl, pH 6.8; 0.1% TEMED; 0.5% APS) which covered the running gel and held the comb.

Staining of proteins on a PAGE gel with Coomassie Blue

All reagents were prepared on the day of use. After electrophoresis, the acrylamide gel was soaked in a gel-fixing solution for 1 hour (50% ethanol and 10% acetic acid in dH_2O). The solution was then aspirated from the gel using a pipette, and enough gel-washing solution (50% methanol, 10% acetic acid in dH_2O) was added to cover the gel completely. The gel was incubated in the gel-washing solution for 4 hours on a rotating incubator. Following the incubation, the gel-washing solution was aspirated from the gel, and the gel was covered with 400ml Coomassie stain (0.1% (w/v) Coomassie blue R350, 20% (v/v) methanol, and 10% (v/v) acetic acid) and stained with gentle agitation for 3hours, after which time the stain was removed by aspiration. The gel was then covered with 300ml destain solution (50% methanol, 10% acetic acid in dH_2O), and incubated with gentle agitation. The extent of destaining was monitored throughout the incubation period, and

the destain solution was changed every 30 minutes. Once the protein bands could be seen without a strong background staining, the gel was equilibrated in 5% acetic acid for at least 1 hour.

Transfer of protein to a nitrocellulose membrane (Western blot)

After electrophoresis, or after Coomassie staining, proteins were transferred from the acrylamide gel to a soaked Protran nitrocellulose membrane using a Trans-Blot electrophoretic transfer cell. The membrane and gel were submerged in blotting buffer (25mM Tris and 192mM glycine in 20% ethanol) and blotted overnight at 150mA at 4°C.

Blocking the membrane and binding of antibodies

After transfer of proteins, the membrane was blocked in blocking buffer (0.2% Tween20 and 2% Marvel milk powder in PBS) for 2 hours, followed by a 2 hour incubation with a primary antibody (1:1000 dilution). The milk protein from the blocking buffer prevented non-specific binding of the primary antibody to the membrane. The primary antibody used was a sheep-anti-Drosophila-GS polyclonal antibody, received as a gift from Lesley McLellan (Dundee University). After incubation with the primary antibody, the membrane was washed four times for 10 minutes in blocking buffer. The membrane was then incubated with the secondary antibody (1:15000 dilution) for 1 hour, after which time the membrane was washed in blocking buffer three times for 10 minutes, followed by a final wash in PBS for 10 minutes. The secondary antibody used was a donkey anti- sheep peroxidase conjugated antibody.

Enhanced chemiluminescence (ECL) detection

The Western Lightning Chemiluminescence Reagent was used to visualise the peroxidase label on the secondary antibody. Peroxidase catalyses the chemiluminescent reaction where light is emitted from the oxidation of luminol. 5ml of the oxidising reagent and 5ml of luminol reagent are added together immediately before addition to the washed membrane. The membrane was incubated in this mixture for exactly 2 minutes before blotting it dry, and visualising on the Typhoon scanner in Chemiluminescence mode.

2.21 S2 cell culture

Drosophila Schneider2 (S2) cells (Schneider, 1972) were obtained as a gift from Carol Midgley (Open University). All work with the cells was performed in a laminar flow hood using sterile technique.

Preparation of SFM

Serum-Free Medium (SFM) was supplemented with 18mM L-glutamine, 50U/ml penicillin and 50ug/ml streptomycin. Once the supplements were added to the SFM, the media was stored at 4°C.

Thawing S2 cells from a frozen stock

A frozen stock of S2 cells was kept in liquid nitrogen storage. When required, a 1ml vial of cells was thawed, and then resuspended in 5ml SFM. The cells were centrifuged at 1000g, and the medium containing DMSO from the frozen cells was discarded. The pellet was resuspended in 10ml fresh SFM, and grown in a 75cm^2 tissue culture flask at 27°C in a normal atmosphere until the cells reached a density of ~10 x 10^6 cells/ml.

Routine maintenance of S2 cells

Cells were maintained by splitting at a 1:10 dilution into SFM every 5-7 days, retaining 1/10 of the conditioned media. The cells were maintained in 75cm^2 flasks at 27°C in a normal atmosphere. S2 cells grow both in suspension and attach lightly to the flask, so in order to remove cells attached to the flask, the flask was tapped, and growth media was aspirated over the attached cells to dislodge them. Most cells would be in suspension after this treatment, and could be used immediately for centrifugation or dilution.

2.22 Cell viability assay and cell counting

Viability of cells was determined by staining with 0.4% Trypan Blue and counting the viable cells using an improved Neubauer haemocytometer. Trypan blue stains dead cells blue, while viable cells exclude the dye and remain clear. To count the number of cells in a solution, an equal volume of cells was added to Trypan blue solution in an eppendorf tube,

and mixed by gently pipetting up and down. 10µl of this solution was then immediately added under the coverslip placed on the haemocytometer, using capillary action, making sure that both sides of the counting chamber were not over-filled or under-filled. After 1 minute, cells were counted by counting the number of cells in the four outer $1mm^2$ grids using an inverted microscope. At least 100 cells were counted in each grid. Living and dead cells were counted separately to allow for the calculation of percentage viability. The number of cells/ml was calculated by multiplying the average number of cells counted in each $1mm^2$ grid by the dilution factor multiplied by 10^4, ie:

Cells/ml = (total number of cells counted in four $1mm^2$ blocks)/4 x 2 x 10^4

Viability was determined as the ratio of the percentage of live cells to the total number of cells.

2.23 Double stranded RNA for RNAi

Double stranded RNA (dsRNA) was made using a Ambion's MEGAscript RNAi kit. A primary DNA template containing opposing T7 RNA polymerase binding sites (GAATTAATACGACTCACTATAGGGAGA) at each end was made by PCR using primers specific to GS, GCLC or GFP, obtained from the MRC's Geneservice RNAi construct library (http://www.geneservice.co.uk/products/rnai/Dros_RNAi.jsp) in the case of GS, or designed from the cDNA sequence in the case of GCLC and GFP. Geneservice is a collection of RNAi constructs which can be purchased. The constructs were created by amplifying *Drosophila* genomic/cDNA fragments using gene-specific primers with dual T7 promoter sequences, and re-amplifying the PCR products to produce DNA templates which are used to synthesise dsRNA.

The primary PCR products were purified using QIAquick PCR purification kit, and quantified using OD.

Using the MEGAscript RNAi kit, two complementary RNA transcripts were synthesised from the DNA template by T7 RNA polymerase and were hybridised together to form a

dsRNA product. DNA and single-stranded RNA was then digested by nucleases before performing the final dsRNA purification with a solid-phase absorption system. The Ambion protocol was followed for a 20µl reaction: 1µg of the appropriate PCR product, 2µl 10X T7 reaction buffer, 2µl of each of the dNTP solutions and 2µl T7 enzyme mix was incubated together at 37°C overnight. The mixture was then incubated at 75°C for 10 minutes, followed by a slow cooling to room temperature to allow the formation of RNA duplexes.

DNA and ssRNA were then removed by nuclease digestion as follows: 20µl of the dsRNA from previous step was added to 5µl digestion buffer, 2µl DNase I and 2µl RNase in a 50µl reaction and incubated at 37°C for 1 hour. After the nuclease digestion, the remaining dsRNA was purified from proteins, free nucleotides, and nucleic acid degradation products in a filter column system. The amount of dsRNA recovered was usually ~1-2µg RNA/µl.

2.24 RNAi in S2 cells

S2 cells were diluted to a density of 3 x 10^6 cells/ml in SFM. The cells were assayed using Trypan Blue to ensure >90% were alive before continuing. The cells were transfected with dsRNA and GeneJuice, a cationic lipid transfection agent, using an adaptation of the manufacturer's protocol:

1ml of a cell suspension at a density of 3 x 10^6 cells was plated per well of a 6-well cell suspension culture dish and the cells were allowed to settle overnight. The following day 2µg dsRNA in 100µl SFM was complexed with 10µl transfection agent in 100µl SFM. This complex was incubated at room temperature for 15 minutes and then added to cells in the wells. The plates were swirled vigorously, followed by a 4 hour incubation at room temperature in the laminar flow hood. A second addition of dsRNA with the transfection agent, along with a further 1.5ml SFM, was pipetted into the the cell/lipoplex mixture in the wells, to bring the total volume in each well to 3ml. The cells were then incubated for

24 hours at 27°C to allow for turnover of the target protein. After the 24 hour incubation, cells were ready for use in further experiments.

2.25 Experimentally induced stress and S2 cell viability

S2 cells treated with RNAi and control cells treated with the lipid transfection agent only were grown at 27°C in SFM for 24 hours, as described above. The following day cells were exposed to DEM or H_2O_2.

2.25.1 DEM treatment

Two concentrations of DEM were tested on cells: a low dose of 0.25mM, and a high dose of 0.5mM. A stock solution of 300mM was created by adding 10µl of a 6M stock to 190µl DMSO. Either 5µl or 2.5µl of this stock solution was added to the 3ml of cells in each well, depending on whether the high dose or low dose of DEM was required. After addition of the DEM, the plates containing the cells and DEM were swirled gently to mix, followed by a 24 hour incubation at 27°C. Following the incubation, cells could be assayed at the first time point of Day 1. Day 0 was used to indicate samples taken directly after DEM addition.

2.25.2 H_2O_2 treatment

Two concentrations of H_2O_2 were tested on cells: a low dose of 10mM, and a high dose of 30mM. A stock solution of 30% H_2O_2 was purchased from Sigma. Either 10.2µl or 3.4µl of this stock solution was added to the 3ml of cells in each well, depending on whether the high dose or low dose of H_2O_2 was required. After addition of the H_2O_2, the plates containing the cells and DEM were swirled gently to allow for mixing. T0 indicated the samples taken directly after the addition of H_2O_2. The assay used for testing cell viability, and consequently GS and GCLC expression levels in cells exposed to H_2O_2 required samples to be taken every 2 hours over a total assay period of 8 hours. The assay was modified from the peroxide viability assay on S2 cells from Radyuk et al (2003).

2.26 Comet assay

The comet assay, also known as the single cell gel assay, is used for evaluating DNA damage in cells. Denatured, damaged DNA will migrate out of a cell under an electric field while undamaged, intact DNA remains in the nucleus and migrates much slower under electrophoresis. The shape and migration pattern of the 'comet tail' is used to assess the amount of DNA damage. The protocol used in this study is a combination of various techniques, which follow the same principle. Cells were embedded in low melting point agarose on a glass slide, then lysed and treated with an alkaline solution to denature the DNA and hydrolyse sites of damage. The cells were then electrophoresed and stained with SYBR green dye. The samples were then visualised by epiflourescence microscopy on an Olympus BX61 microscope using the FITC filter. Analysis was carried out using CometScore version 1.5 Freeware, available from http://autocomet.com, using 100 randomly selected images. The full protocol and reagents used are detailed below:

10^5 cells were added to 75µl 0.5% low melting point agar solution made up in PBS, and warmed to 37°C. This mixture was spread over a frosted glass microscope slide in duplicates and left to cool at 4°C for 10 minutes. The slide was then placed in pre-chilled lysis solution (2.5M NaCl, 100mM EDTA, 10mM Trizma base) at 4°C for 1 hour. After 3 washes in a 0.4M Tris pH 7.5 solution, the slides were placed in an alkaline solution (300mM NaOH, 1mM EDTA, pH12) for 30 minutes. Samples were rinsed in 1X TBE three times, and then transferred to a horizontal electrophoresis apparatus and covered with 1X TBE. Power was set to 1V per cm, and voltage was applied for 10 minutes. Slides were then dipped in 70% ethanol for 5 minutes, followed by air drying. At this stage, slides could either be stored with desiccant for later use, or stained immediately with 50µl diluted SYBR Green solution (1µl 10 000X SYBR Green in 10ml TE buffer). Analysis of the images is described in Chapter 7.

2.27 Embryo micro-injection to produce transgenic flies

Human-GS (hGS) cloned into pUAST vector was injected into *Drosophila* embryos in an attempt to create a transgenic fly which could over-express hGS. The embryos used were from a fly strain that expressed a transposase, allowing chromosomal integration of the transgene flanked with a P-element sequence (Spradling and Rubin, 1982). Mobilisation of the transgene from the plasmid onto a chromosome allowed for transgene integration into the germline.

DNA used for injecting into embryos was purified by ethanol precipitation to remove any salts and impurities. The DNA was resuspended in 1X injection buffer (10X buffer: 0.1M Sodium Phosphate Buffer pH 6.8, 5mM KCl, filter through 0.22μm syringe filters) to give a final concentration of 1μg/μl. Prior to loading the DNA into the injection needle, the DNA solution was centrifuged at 13 000rpm in a microfuge for 5 minutes to pellet any undissolved DNA.

In order to collect embryos at a developmental stage prior to blastoderm cellularisation, fertilised female flies were allowed to lay eggs in a laying cage on agar plates overnight in a 25°C incubator. In the morning, a fresh injection plate was placed into the cage for 1 hour. After this time period, eggs were collected from the plate, and the procedure of putting in fresh plates and collecting embryos was repeated throughout the day at hourly intervals. The pre-blastoderm stage required for injection occurs approximately 40 minutes after egg laying.

The embryos were dechorionated by rolling them carefully on double sided sticky tape until the outer shell has been removed. Between 20 and 50 dechorionated embryos were lined up on double-sided sticky tape on the edge of a glass slide with the posterior end of the embryo facing the closest edge of the slide, in a process not exceeding 15 minutes. The embryos were placed in a desiccator for 5 minutes before being covered with Halocarbon oil to prevent further desiccation. Eggs were typically collected every 30-60 minutes at

25°C and injected within 30 minutes of collection to attempt to prevent an abundance of embryos which had passed Bownes stage 2, after which cellularisation occurs.

A glass micro-injection needle was used to inject purified DNA into the embryos which had not yet passed to Bownes stage 2 (pre-blastoderm). 2µl of the centrifuged DNA solution (1µg/µl) was pipetted into the needle. The needle was then attached to the injection apparatus, which consisted of an inverted microscope (Leica), a micromanipulator, and an air-pressure injecting device connected to a needle holder (Micro Instruments). The slide containing the lined-up embryos with the posterior sides facing the needle was placed on the microscope stage, and the needle containing the purified plasmid was brought into the same focal plane as the embryos. The needle was brought down slowly into the oil, and gently pressed further to break the tip of the needle. The jagged edge of the needle facilitates penetration of the embryo. The needle was brought close to the posterior side of the first embryo, and the injection button attached to the needle was pressed slowly to allow the release of a just-visible amount of DNA into the cell. The procedure was repeated for the whole row of embryos.

After injection, the embryos were carefully placed into a shallow hollow made in standard fly food in a vial. The embryos were then allowed to develop at 25°C, keeping a close check on the developmental stage. At the pupal stage, each pupa was moved to a separate vial and allowed to develop to hatching. The resultant white eyed adults (F_0) were crossed to w^{1118} adults. Orange eyed F_1 progeny in a w^{1118} background indicated the presence of the transgene marked with w^+. The transgenic flies were then collected and crossed to make homozygous stocks.

2.28 Oxidative stress tests on flies

Two to three day old flies were collected for the oxidative stress assay, and placed in a transparent polyethylene tub with a lid (Figure 2.2) at a density of approximately 75 flies per tub. After a 48 hour aging period with access to regular fly food (Figure 2.2A), flies

were starved for 4 hours before the addition of a 10% sucrose solution containing either 5mM Paraquat or 6mM DEM to absorbent cotton wool in the fly tub (Figure 2.2B). The DEM or Paraquat solution was added to the cotton wool every second day, allowing a constant supply of the chemical stressor. The experiment was run at 25°C in the dark, because Paraquat is light sensitive (Wang et al., 2000). Dead flies were counted every day, and the data were analysed using the SPSS statistical software application (2004, v13.0). Each experiment was completed at least in duplicate, and usually in triplicate. The data from all three experiments were pooled to increase the power of the statistical tests. Mean and maximal survival was calculated using the Kaplan-Meier analysis. The Mann-Whitney non-parametric test was used to determine if there was a significant difference between two groups.

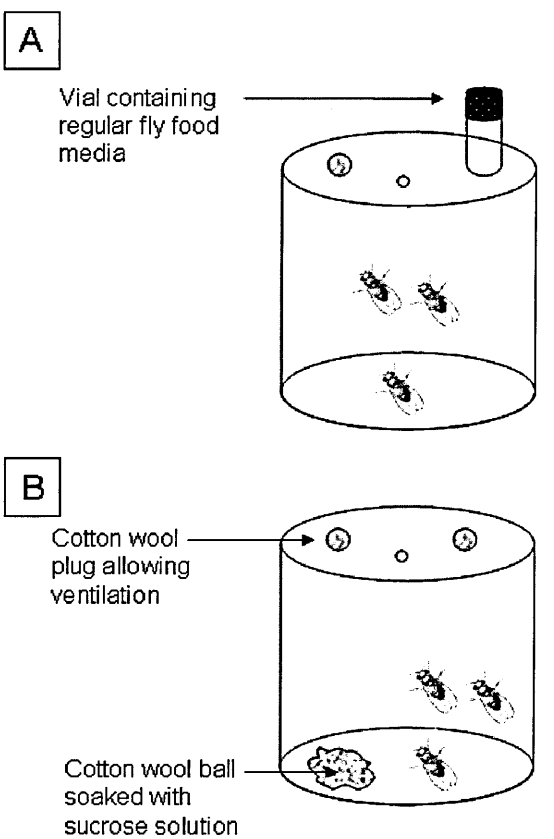


Figure 2.2: Survival tubs used for oxidative stress experiments. Panel A shows the tub during the 2 day ageing period where flies are fed on regular fly food. Panel B shows the tub during the administration of chemical oxidative stress. The vial containing food is removed, and cotton wool soaked in a sucrose solution containing the appropriate chemical stressor is introduced to the tub.

2.29 RU486 administration with the geneswitch system

The Geneswitch system present in the P{ELAV-Geneswitch}stock (Osterwalder et al., 2001; Roman et al., 2001) uses a GAL4-RU486 receptor fusion protein to allow conditional GAL4 expression in specific tissues controlled by the presence or absence of the progesterone chemical RU486 (Mifepristone). In the Geneswitch system, GAL4 is expressed in a tissue-specific manner through the use of a specific promoter, as in standard GAL4/UAS systems, however the GAL4 molecule is biologically inactive until it is bound to RU486. The GAL4-RU486 compound is then able to bind to a UAS receptor and thereby activate transcription of the downstream gene.

A 20mM stock solution of RU486 made in 80% ethanol was added to standard fly media to a final concentration of 0.5mM. Flies or larvae were then added to the media containing the RU486 depending on when transcriptional activation of the target gene was required. Transcription is thought to be activated 3 hours after ingestion of RU486, and reaches a maximum level 24hours after ingestion (Roman et al., 2001), therefore flies were tested for gene expression levels at least 24 hours after being transferred to the RU486 containing medium. RU486 has not been found to have any effect on longevity, fecundity or behaviour in *Drosophila* (Roman et al., 2001; Poirier and Seroude, 2005) .

2.30 Fly stocks

Table 2.4: Fly stocks used. All stocks are available from Bloomingtons stock centre, with the exception of the first two stocks which were marked in our laboratory.

Genotype	Stock name (if different to genotype)	Comments
w sn^{3} EP1322	*EP1322*	P-element in the 5′ end of copy1 of GS
w^{1118} sn EP0970	*EP0970*	P-element downstream of GS
w^{1118}		Our lab reference strain
CantonS		Wild type strain
winscy; dp^{ov1}; rho^{ve-1}; sv$^{spa-pol}$		X-balancer, viable
Df(1)C128/FM6		X-balancer chromosome
C(1)DX y f / FM6	*FM6*	X-balancer chromosome
l(1)7Ed1/FM7a	*FM7*	X-balancer chromosome
w sn^{3}		GS single copy reference strain
y sn^{3} v		GS double copy reference strain

Genotype	Stock name (if different to genotype)	Comments
w m f f		
y^2 m^{74F} wy^{74i} sd^1 os^s		X chromosome marker stock
Drosophila mauritiana C164.1		Sibling species
Drosophila simulans Oxnard		Sibling species
Drosophila simulans Ethiopie 225.1		Sibling species
Drosophila simulans C167.4		Sibling species
Drosophila sechellia Jallons		Sibling species
Drosophila sechellia Cousin		Sibling species
Drosophila melanogaster VAG1		Wild type strain
Drosophila melanogaster BER1		Wild type strain
Drosophila melanogaster RC1		Wild type strain
Drosophila melanogaster BS1		Wild type strain
Drosophila melanogaster PRY2		Wild type strain
Drosophila melanogaster Wild 10D		Wild type strain
Drosophila melanogaster CA1		Wild type strain
w sn^3 *l(1)8.3/FM6*	*l(1)8.3/FM6*	GS lethal mutation created by imprecise mutagenesis
w^*; ry^{506} Sb^1 *P99B/TM6B,* Tb^1	Δ2-3	Source of transposase
w^{1118}; *CyO/Bl; TM6,Tb/TM2,Ubx P{ELAV-Geneswitch}*	Elav_GS_GAL4 or GeneSwitch	RU486 dependent GAL4 driver - Neuronal expression
y^1 w^*; *P{*w^{+mC}*=Act5C-GAL4}25FO1/CyO,* y^+	Actin(2)-GAL4	GAL4 driver – ubiquitous expression
y^1 w^*; *P{*w^{+mC}*=Act5C-GAL4}17bFO1/TM6B,* Tb^1	Actin(3)-GAL4	GAL4 driver – ubiquitous expression
w^-; *P{*w^+*, GAL4}-D42*	D42-GAL4	GAL4 driver – motorneuron expression
y^1 w^*; *P{*w^{+mC}*=tubP-GAL4}LL7/TM3,* Sb^1	Tubulin-GAL4	GAL4 driver – ubiquitous expression
w^*;*P{*w^{+mc}*=UAS-GFP.565T}T2*		Ubiquitous UAS-GFP transgene used to test GAL4 drivers (2^{nd} chromosome)
w^*;*P{*w^{+mc}*=UAS-GFP.565T}T2*		Nuclear localised UAS-GFP transgene used to test GAL4 drivers (2^{nd} chromosome)
w^*;*P{*w^{+mC}*=GAL4-Hsp70.PB}2/CyO*	GAL4-HS	GAL4 fused to a leaky heat shock promoter
w^{1118}; *pUAST-Gclc*	UAS-*Gclc*	Contains *Drosophila* GCLC transgene in pUAST vector
w^{1118}; *pUAST-Gclc, pUAST-Gclm*	*Gclc/Gclm*-pUAST	Contains both GCL subunits in pUAST vectors
w^{1118}; *hGS-pUAST (A)*		Contains human GS transgene in pUAST vector inserted onto the second chromosome

Genotype	Stock name (if different to genotype)	Comments
Dp(1;Y) W73, y^{31d} $B^{1}B^{s}$ / C(1)Dx $y^{1}f^{1}/y^{1}$ baz^{EH171}		Marker stock used to map *l(1)8.3* lethality

2.31 Statistical tests

Standard statistical tests including the independent t-test and calculation of the standard error of the means (SEM) were carried out using the Microsoft Office Excel spreadsheet software package (2003, v5.1.2600). Before performing a t-test, the data were tested for normal distribution using the 'descriptive statistics' option on the SPSS statistical software application (2004, v13.0). The t-test was used to compare the means of two different groups of parametric data to determine whether there was a significant difference between the groups or not. Statistical significance was defined as significant when $p<0.05$ (*), and highly significant when $p<0.01$ (**) or $p<0.001$ (***). SEM measures the standard error of the sampling distribution of the means, ie, SEM determines whether the sample mean is a true reflection of the population mean. The equation used to calculate SEM was to divide the standard deviation of the mean by the square root of the number of observations.

In the case of non-parametric data (non-normal distribution), such as fly survival data, the Mann-Whitney U test was used to test for differences between two groups using the SPSS software application. The significance level (labelled 'Asymp.Sig.' by SPSS) indicated whether there was a statistically significant difference between groups or not, where values less than 0.05 (*) defined a statistically significant difference, and values less than 0.01 (**) or less than 0.001 (***) were highly significant.

2.32 Details of Suppliers and Catalogue number of routinely used reagents

Reagent	Supplier	Catalogue number
1Kb plus DNA ladder	Invitrogen	10787-018
Agar	Sigma	A5054
Agarose	Invitrogen	15510-027

Reagent	Supplier	Catalogue number
Agarose (low melting point)	Sigma	A9414-5G
Ammonium persulphate (APS)	Biorad	161-0700
Ampicillin	Sigma	A9393
Anti-sheep peroxidase conjugated antibody	Sigma	A3415
β-mercaptoethanol	Sigma	M7154
Benzamidine	Sigma	12072
Bovine serum albumin (BSA)	Sigma	A2153
Bromophenol blue	Sigma	B0126
Chloramphenicol	Sigma	23275
Chloroform	Sigma	C2432
Coomassie Blue solution	BioRad	161-0436
dCTP-^{32}P	Perkin Elmer	BLU013H
dNTP solution set	NEBiolabs	N0446S
DEM (diethyl maleate)	Sigma	D97703
DEPC	Sigma	D5758
N,N-dimethylformamide	Sigma	D4551
Disodium hydrogen phosphate	Sigma	S3264
Dithiothreitol (DTT)	Sigma	D0632
Drosophila Serum Free Medium (SFM)	Invitrogen	10797-017
DTNB	Sigma	D8130
EDTA (sodium salt)	Sigma	E6760
Electroblotting buffer (western blots)	National diagnostics	EC880
EMM -Leu medium	Formedium	PDM0810
Ethidium bromide	Sigma	E1510
Ficoll 400	Sigma	F2637
Formamide	Sigma	PS2031
Glacial acetic acid	Sigma	537020
Glucose	Sigma	D8066
L-Glutamine 100X	Invitrogen	25030-024
Glutathione (reduced)	Sigma	G6529
Glycerol	Sigma	G5516
Glycine	Sigma	G8898
Halocarbon oil 700	Sigma	H8898
Hydrochloric acid	Fluka (Sigma)	08256
Hydrogen peroxide (30%)	Sigma	H1009
Hyperladder II molecular weight marker	Bioline	BIO-33039
Isopropanol	Sigma	I9516
Insect GeneJuice	Novagen	71259-3
Lithium chloride	Sigma	L4158
Magnesium chloride	Sigma	M8266
Magnesium sulphate	Sigma	M2643
Mifepristone (RU486)	Sigma	M8046
MOPS	Sigma	M1294
NADPH	Sigma	N1630
Oligo(dt)$_{12-18}$ Primer	Invitrogen	18418-012

Reagent	Supplier	Catalogue number
Parafilm M	Sigma	P7793
Paraquat (methyl viologen)	Sigma	M2254
PEG 4000	BDH Laboratory supplies	29576
Penicillin/Streptomycin 100X	Invitrogen	15070-063
Phenol - chloroform - isoamyl alcohol (25:24:1)	Sigma	77617
PMSF	Sigma	P7626
Polyvinylpyrrolidone	Sigma	P5288
p-hydroxymethyl benzoic ester	Sigma	H6654
Potassium acetate	Sigma	P5708
Potassium chloride	Sigma	P9541
Protogel acrylamide	Geneflow	A2-0074
Rainbow molecular weight marker	Amersham	RPN756
RNA ladder, 0.24 -9kb	Gibco-BRL	15620-016
Running buffer 10X (SDS-PAGE gels)	National diagnostics	EC870
Sodium chloride	Sigma	S3014
Sodium dihydrogen phosphate	Sigma	S3139
Sodium dodecyl sulphate (SDS - 20%)	Sigma	05030
Sodium hydroxide	Sigma	480879
Sodium phosphate bibasic pentahydrate	Sigma	04283
Sulphosalycyclic acid	Sigma	S0640
SYBR green	Sigma	S9430
TEMED (N,N,N′,N′-Tetramethylethylenediamine)	Sigma	87689
Trichloroacetic acid (TCA)	Fluka (Sigma)	91228
Tri-sodium citrate	Sigma	S1894
Trizma base	Sigma	T1503
TRIzol	GibcoBRL	15596-018
Tris-HCl	Sigma	T3253
Triton X-100	Sigma	234729
Trypan Blue	Sigma	T8154
Tryptone	Sigma	T9410
Tween20	Sigma	P7949
Urea	Sigma	U5378
Water (tissue culture grade)	Sigma	W3500
X-gal	Sigma	B6024
Xylene cyanol	Sigma	335940
Yeast extract	BD	288620

Kit & enzymes & vectors	Supplier	Catalogue number
Antarctic phosphatase	NEBiolabs	M0289S

BioRad DC protein assay	BioRad	500-0116
Calf Intestinal Phosphatase	NEBiolabs	M0290S
Glutathione reductase	Sigma	G3664
Megaprime DNA labelling system	Amersham	RPN1606
MEGAscript RNAi kit	Ambion	1626
One Shot Top 10 chemically competent cells	Invitrogen	C4040-10
Pfu ultra Taq	Stratagene	600380
Plasmid Maxi Kit	Qiagen	12163
pNMT1-TOPO	Invitrogen	K181-01
SpinPrep gel DNA kit	Novagen	70958-3
SuperScript II Reverse Transcriptase	Invitrogen	18064-022
T4 DNA Ligase	Invitrogen	15224-017
Taq DNA Polymerase with 10X buffer	Sigma	D1806
TOPO TA cloning kit	Invitrogen	K4500-01
QIAquick PCR purification kit	Qiagen	28104
Quick ligation kit	NEBiolabs	M2200S
Western Lightning Chemiluminescence Reagent *plus*	Perkin Elmer	NEL103
Wizard Plus SV Minipreps DNA purification system	Promega	A1330
XL1 Blue *E.coli* cells	Stratagene	200130

Equipment/consumables	**Supplier**	**Catalogue number**
0.22 micron syringe filters	Sigma	F8148
0.45 micron syringe filters	Sigma	F8273
3MM blotting paper	Whatman	3030-866
6-well cell suspension tissue culture dish	Greiner	657-185
DNA Engine Opticon2 thermocycler	MJ Research	n/a
Disposable pellet pestle	Sigma	Z35-9947
Electroporator	Stratagene	n/a
Electroporation cuvette	Flowgen	C24182
Fly bottles	Scientific Laboratory	INC 9002
Fly vials	Sarstedt Ltd	58-490
Haemocytometer	Scientific Laboratory Supplies	HAE2112
Hybond N^+ nylon membrane	Amersham	RPN203B
Hybond N nylon membrane	Amersham	RPN203N
Low profile 0.2ml PCR tubes	MJ Research	TLS-0801
Flat caps for 0.2ml PCR tubes	MJ Research	TCS-0803
Microfuge	Fisher Scientific	CFA-105-010H
mini-PROTEAN II system	Biorad	165-2944
Non-absorbent cotton wool	Richardsons	C1690

Equipment/consumables	Supplier	Catalogue number
Orbital shaker	Stuart Scientific model SO1	n/a
Petri dishes (90mm)	Greiner	633185
Pipette tips 1-10µl	Starlabs	S111-3700
Pipette tips 1-200µl	Starlabs	S1110-1800
Pipette tips 100-1000µl	Starlabs	S1111-2721
Pipette tips (filtered) 1-20µl	Starlabs	S1120-1810
Pipette tips (filtered) 1-200µl	Starlabs	S1120-8810
Pipette tips (filtered) 100-1000µl	Starlabs	S1126-7810
Plate reader	Labsystems Multiscan Plus	n/a
Polyethylene tubs and lids for fly survival experiments	Medfor Products Ltd	PN92
Protran nitrocellulose membrane	Schleicher & Schuell	BA83
Rotating hybridisation incubator	Hybaid Micro4	n/a
Sterile falcon tubes	Greiner	188271
Sterile centrifuge tubes (50ml)	Greiner	227270
Sterile pipette, single wrap (5ml)	Greiner	606180
Sterile pipette, single wrap (10ml)	Greiner	607180
Sterile pipette, single wrap (25ml)	Greiner	760180
Tissue culture flask - 75cm^2	Greiner	685-190
Trans-Blot cell	Biorad	170-3939
Typhoon 9410 variable mode imager & ImageQuant software	Amersham biosciences	63-0055-80
UV cross linker	Hoeffer UCV500	n/a
X-Ray film	Scientific Laboratory	MOL 7016

3 Chapter 3: Genomic organisation of *GS* in *Drosophila melanogaster*

3.1 Introduction

Glutathione synthetase (GS) which catalyses the second step of the two step glutathione synthesis reaction is a member of the ATP-grasp family of enzymes (Galperin and Koonin, 1997). This superfamily is characterised by the presence of phosphate binding loop and an ATP binding cleft formed by two structural elements, each composed of two β-sheets and a loop (Murzin, 1996; Galperin and Koonin, 1997). GS catalyses the addition of glycine to γ-glutamylcysteine by phosphorylating the C-terminal carboxylate of γ-glutamylcysteine, which forms an acyl phosphate intermediate. This intermediate is bound by glycine, resulting in the formation of glutathione, ADP and an inorganic phosphate (Jez and Cahoon, 2004).

GS is conserved between species that are evolutionarily diverse, and although the sequence homology can vary to a great degree, there are specific residues which are highly conserved. Characterisation of GS has been undertaken in humans (Webb et al., 1995), mice (Shi et al., 1996), rats (Huang et al., 1995), *S. cerevisiae* (Inoue et al., 1998), *S. pombe* (Kim et al., 2003), *Arabidopsis* (Rawlins et al., 1995), Legumes (Frendo et al., 2001) and *E. coli* (Gushima et al., 1984), while in *Drosophila* there has been no published work on GS.

The mechanism of action of glutathione synthetase in different organisms is conserved despite the lack of extensive sequence homology (Figure 3.1) and differences in molecular masses (Jez and Cahoon, 2004). Studying GS in *Drosophila* should therefore give us information on the functioning of GS that will be broadly relevant. Because glutathione is one of the major cellular antioxidants, understanding the gene structure of GS in *Drosophila* (referred to *DmGS* in this thesis) is important to start dissecting the role of GS

and glutathione in the oxidative stress response and, ultimately, ageing.

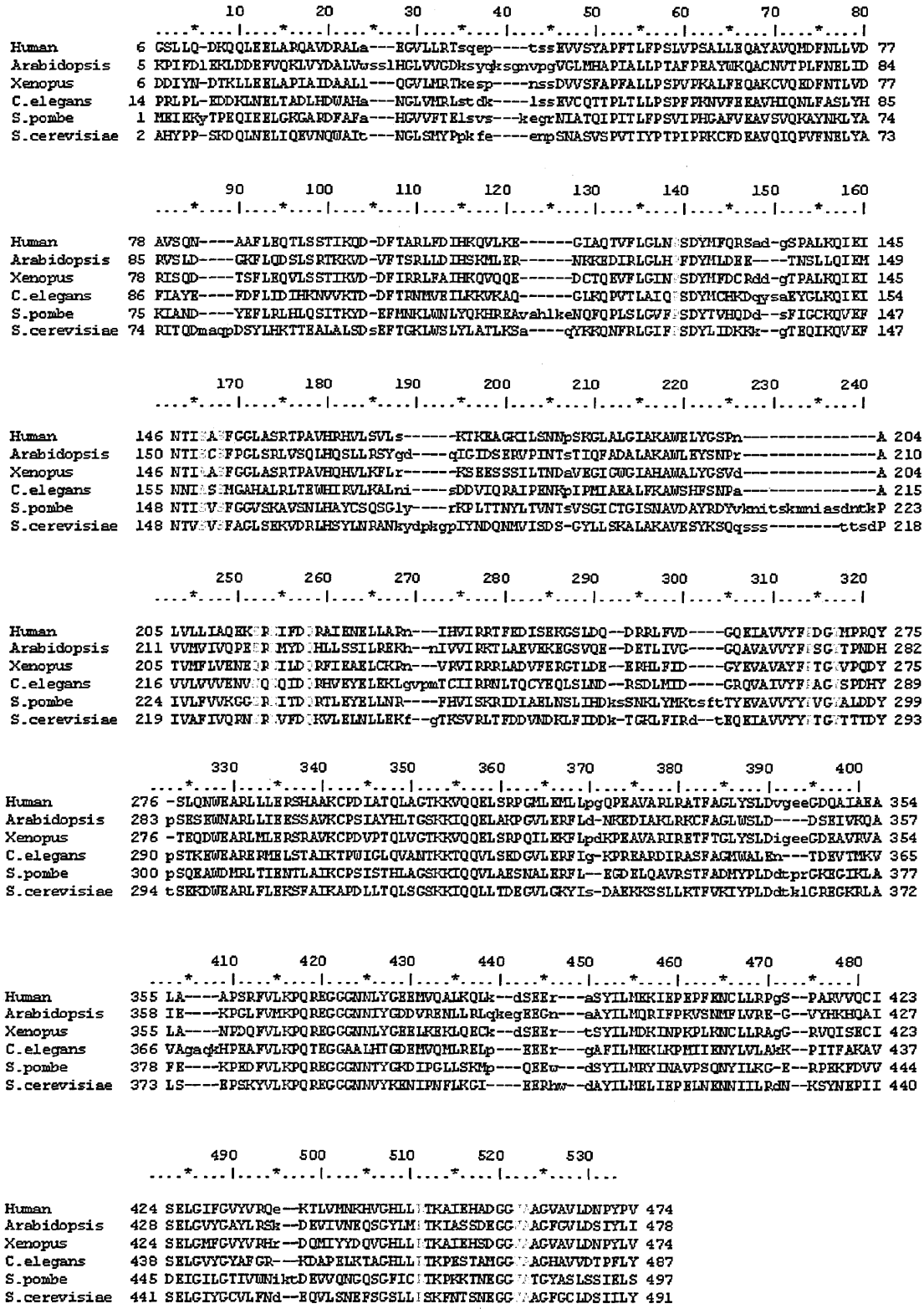


Figure 3.1: Alignment of GS protein sequences from six representative eukaryotic species. Active site residues which form specific binding sites in human GS are labelled in yellow (glutathione binding site), red (ATP binding site) and blue (magnesium binding site). The alignment was created using a Blast search of conserved domains (Marchler-Bauer and Bryant, 2004).

This chapter describes work aimed at characterising the molecular structure of the *DmGS* locus. The findings present an interesting result with the gene structure of *DmGS* appearing to be a duplication. The distribution of this duplication within strains of *Drosophila melanogaster*, as well as in sibling species of *Drosophila,* was studied. Finally, the two gene copies were compared and the differences discussed.

3.2 Results

3.2.1 Molecular characterisation of the GS locus in Drosophila melanogaster

At the outset of the work described in this thesis, the genomic region of the X-chromosome containing *DmGS* had not yet been completely sequenced, making it difficult to establish the structure of the locus using molecular techniques. Preliminary work, discussed further in Chapter 4, was aimed at generating *DmGS* mutants by imprecise excision of an EP-element inserted in the 5′ end of *DmGS* (Kansagra and Saunders, pers. comm.). The reported insertion site was verified by PCR using primers complementary to the 5′ end of the P-element paired with primers complementary to positions downstream of the 3′ end of the P-element, as well as primers pairs complementary to the 3′ end of the P-element and the region upstream of the 5′ end of the P-element (Figure 3.2a). Sequencing the PCR products from primer pairs Plac1 and L3, as well as the product from Pry4 and R1, mapped the P-element to the 5′end of *DmGS* (Kansagra, pers comm.). This corresponds to position 25010 on BACR29H04 (Appendix A), a Bacterial Artificial Chromosome which contains the entire *Drosophila* GS gene region.

Contradictory results were obtained from further PCR analysis of the P insertion in strain EP1322. The primer pairs L3 and Plac1, which correspond to the region upstream of the P-element, and to the 5′ end of the P-element respectively, gave a PCR product of the expected size (500bp), and the primer pairs R1 and Pry4, which correspond to the region downstream of the P-element and the 3′ end of the P-element respectively, also resulted in

a PCR product of the expected size (300bp). An anomaly was noted, however, when primers L3 and R1, which flank the P-element insertion site, yielded a product characteristic of wild type GS from EP1322 genomic DNA (Figure 3.2), even though that region of the gene should have been disrupted by the 8kb P-element insertion whose position had already been confirmed.

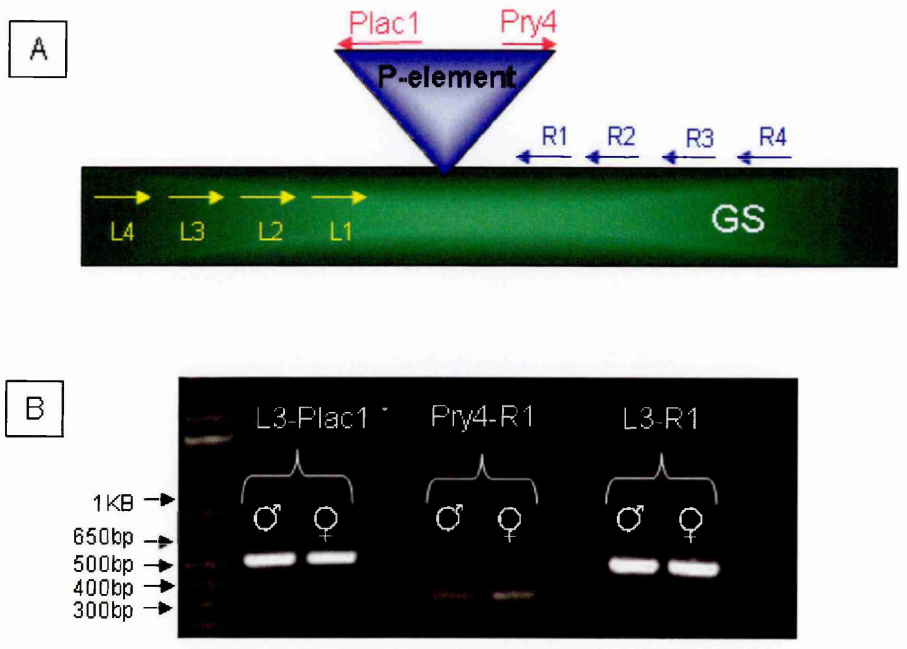


Figure 3.2: Determining the position of the P-element with the *DmGS* locus in strain EP1322. Panel A is a schematic representation of the hypothetical placement of the 8KB P-element in *DmGS*, with the positions of the various primers along *DmGS* shown as arrows 'Left 1-4' and 'Right 1-4'. Primers on the 3′ and 5′ ends of the P-element, Plac1 and Pry4, are shown in pink. Panel B is an agarose gel with PCR samples loaded. PCR was performed on DNA from male as well as female EP1322 flies. The first two lanes are PCR products from the primer pair L3 on *DmGS* and Plac1 on the P-element. The middle two lanes are PCR products from the primer pair Pry4 on the P-element and R1 on *DmGS*. The last two lanes are PCR products from the primer pair L3 and R1.

One of the possible explanations for this unexpected result was the presence of a second copy of the *DmGS* gene which did not contain a P-element. Southern blotting of EP1322 genomic DNA confirmed the presence of a duplication of *DmGS* in strain EP1322, and was strongly suggestive of a tandem duplication. This was corroborated by the then newly published version 3 of the *Drosophila* sequence release (Celniker et al., 2002). After verifying the existence of a *DmGS* duplication, the presence of the *DmGS* duplication in other strains of *D. melanogaster* was tested by PCR.

3.2.1.1 Southern Blot of EP1322 genomic DNA reveals presence of a *DmGS* duplication

A Southern blot was performed on EP1322 genomic DNA to determine whether a duplication of *DmGS* was present in this strain. By digesting the region of the genome containing *DmGS* with the restriction enzyme HaeII, which has recognition sites within as well as flanking the P-element, a unique digestion pattern would be expected for each copy of the *DmGS* gene provided there was a gene duplication. Strains EP0970 and w^{1118}, which did not contain a P-element within *DmGS*, were used as controls.

The probe used in the Southern blot was amplified by PCR from BACR29H04 DNA using the primers L3 and R4, which are complementary to the region spanning positions 24740-26320 on the BAC (Appendix A). The 1580bp probe corresponds to a genomic region extending over the insertion site of the P-element in strain EP1322.

As illustrated in Figure 3.3A, in the presence of the P-element, the probe was expected to bind to 3 fragments of DNA after digestion with HaeII; A (2392bp), B (705bp) and C (116bp). In the absence of the P-element, the probe should bind to two fragments of DNA; D (1081bp) and C (the identical fragment produced in the presence of EP1322).

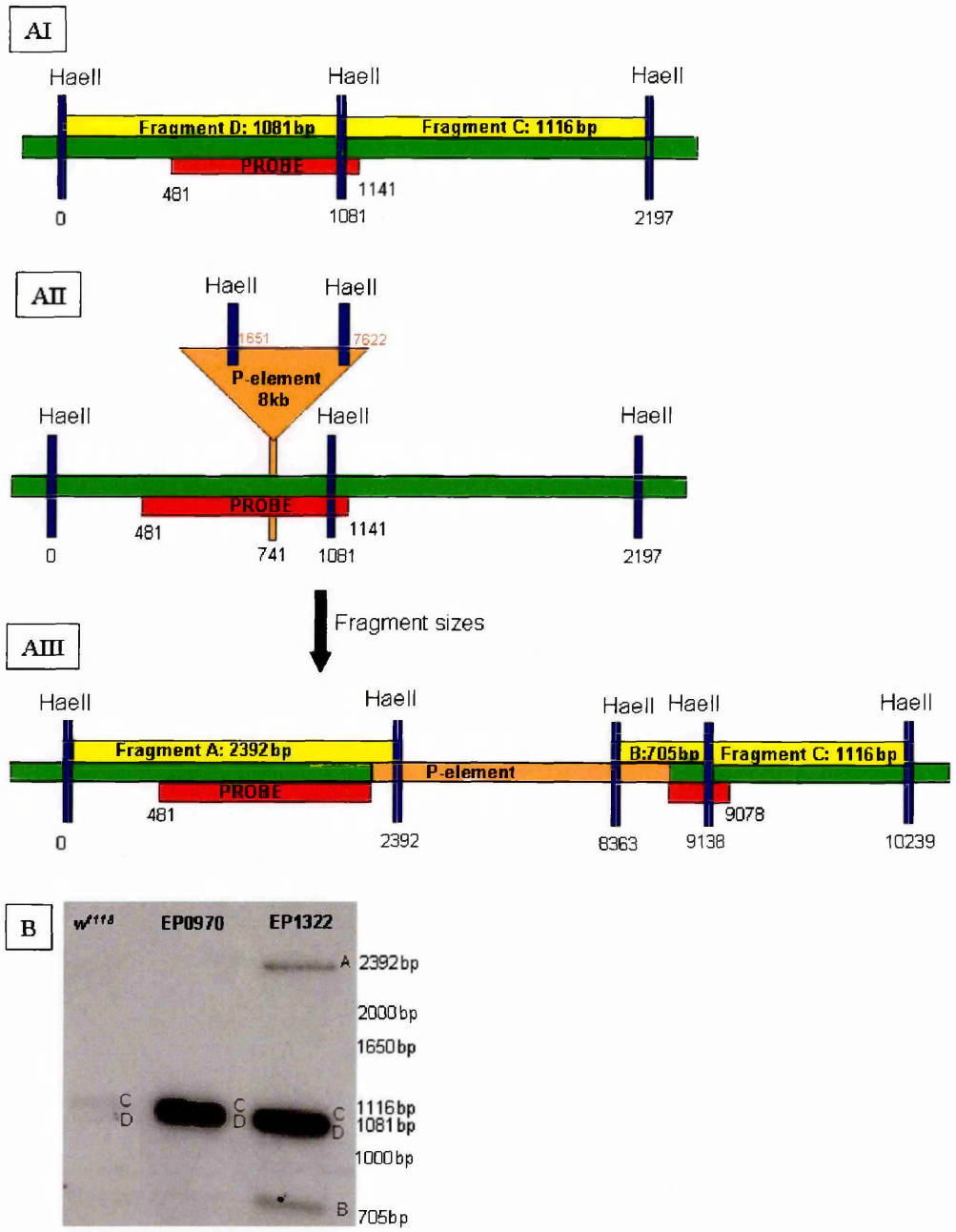


Figure 3.3: Panel A: Schematic representation of the *DmGS* genomic region (green bar) digested with HaeII and probed with L3-R4 probe (red bar). The positions of HaeII restriction sites are marked with blue bars, numbering starts at 0 at the first HaeII site. The position of the probe is also shown underneath the probe on the diagram. Genomic DNA from EP1322 and control male flies was digested with HaeII and in a Southern blot was probed with a PCR fragment spanning the region of the P-element insertion. Panel AI illustrates the standard 5′ *DmGS* genomic region: Fragments D and C (yellow bars) were expected after digestion. Panel AII shows the position of the P-element in the 5′ end of *DmGS*, along with the HaeII restriction sites. Panel AIII illustrates the expected fragment sizes in the case where the P-element is present in the 5′ end of *DmGS*: Fragments A, B and C (yellow bars) were expected after HaeII digestion.
Panel B: Southern Blot results of HaeII digestion of EP1322 genomic DNA from adult males, with control strains EP0970 and w^{1118}, probed with fragment L3-R4. EP1322 DNA has four bands: Fragments C and D expected in the absence of the P-element EP1322, as well as the additional two fragments (A and B) expected only in the presence of P-element EP1322. Bands of 705bp (fragment B), 2392bp (fragment A) and 1116bp (fragment C) correspond to the sizes expected in the presence of the P-element, and the fragments of 1116bp (fragment C) and 1081bp (fragment D) correspond to the sizes expected in the absence of the P-element. The control DNAs had 2 fragments of 1116bp and 1081bp (fragments C and D), which was expected of the genomic region probed after digestion with HaeII.

The Southern blot of EP1322 genomic DNA resulted in bands corresponding to sizes expected in the presence, as well as the absence of the P-element, thus providing proof of two distinct copies of *DmGS* in this strain (Figure 3.3B). The blot results for the controls w^{1118} and EP0970 gave the expected results of two bands of 1081bp and 1116bp as there was no P-element present in *DmGS* in these strains. The two bands of 1081bp and 1116bp in EP1322 and EP0970 are not clearly resolved in this blot since there is only a 35bp difference in size between the two bands. A shorter exposure time would have shown the presence of the two separate bands more distinctly.

3.2.1.2 Bioinformatic evidence of a *DmGS* duplication

Further evidence of the presence of a *DmGS* gene duplication was obtained with release 3 of the *Drosophila* genome sequence, in which sequencing of the region including *DmGS* was completed. A map of the *DmGS* region reveals a tandem duplication of *DmGS* (Figure 3.4A). There is an internally deleted *297* element inserted into the middle of the distal copy of the *DmGS* gene (CG6835). Full length *297* elements are 6.9kb (Dominguez and Albornoz, 1999): the *297* element within CG6835 has an internal deletion of 4.26kb. The function of CG6835 is annotated as glutathione synthetase on FlyBase. Its gene structure appeared very similar to the neighbouring gene CG32495. A sequence alignment of those two genes revealed them to be identical, other than the presence of the internally deleted *297* element within the large intron of CG6835. The presence of a *DmGS* gene duplication has since been confirmed in the latest release (version 4.2.1, September 2005) of the annotated *Drosophila* sequence, which assigned a glutathione synthetase function to GC32495, the proximal (to the centromere) copy of *DmGS* (Drysdale and Crosby, 2005), as shown in Figure 3.4B.

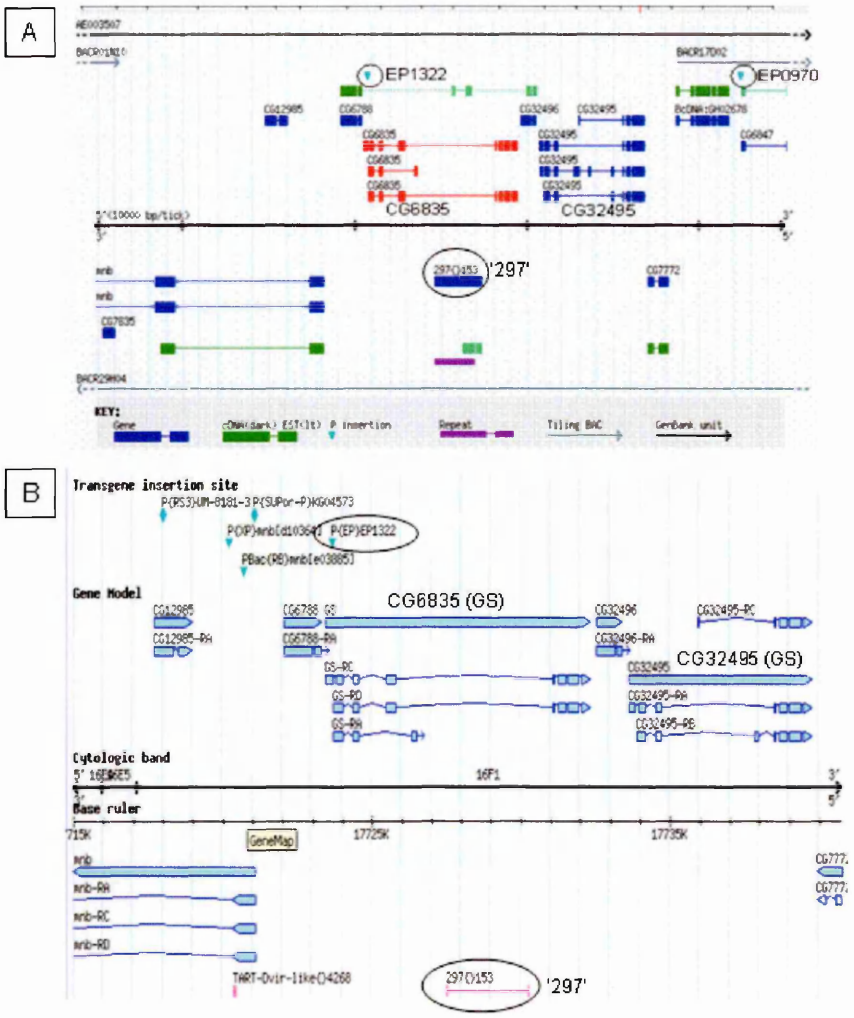


Figure 3.4: Schematic representation of *DmGS* gene structure as illustrated on FlyBase. Panel A is the original version 3 release that was available after the start of this project (Celniker et al., 2002). The green rectangles represent cDNA ESTs, the blue rectangles represent exons of genes, and the turquoise inverted triangles represent P-element insertions. The gene highlighted in red (CG6835) is *DmGS* distal copy, and CG32495 is the proximal copy of *DmGS*. Panel B is the current gene model of the *DmGS* locus from Release 4.2.1 (September 2005) of the annotated sequence on FlyBase (Drysdale and Crosby, 2005). The 297 element insertion site is circled, as is the position of EP1322.

3.2.2 Determination of the presence of the GS duplication within Drosophila strains.

In order to determine how widespread the *GS* duplication is within *D. melanogaster* and its sibling species, a PCR test was carried out on a number of strains. The primer pair GS_Ex12F and GS_R4 was designed in order to amplify a fragment between the last exon of the distal copy to the first exon of the proximal copy of *DmGS*, a specific amplification product of the region between the two *DmGS* copies (Figure 3.5). This PCR reaction allowed a rapid means of detecting the presence or absence of the duplication.

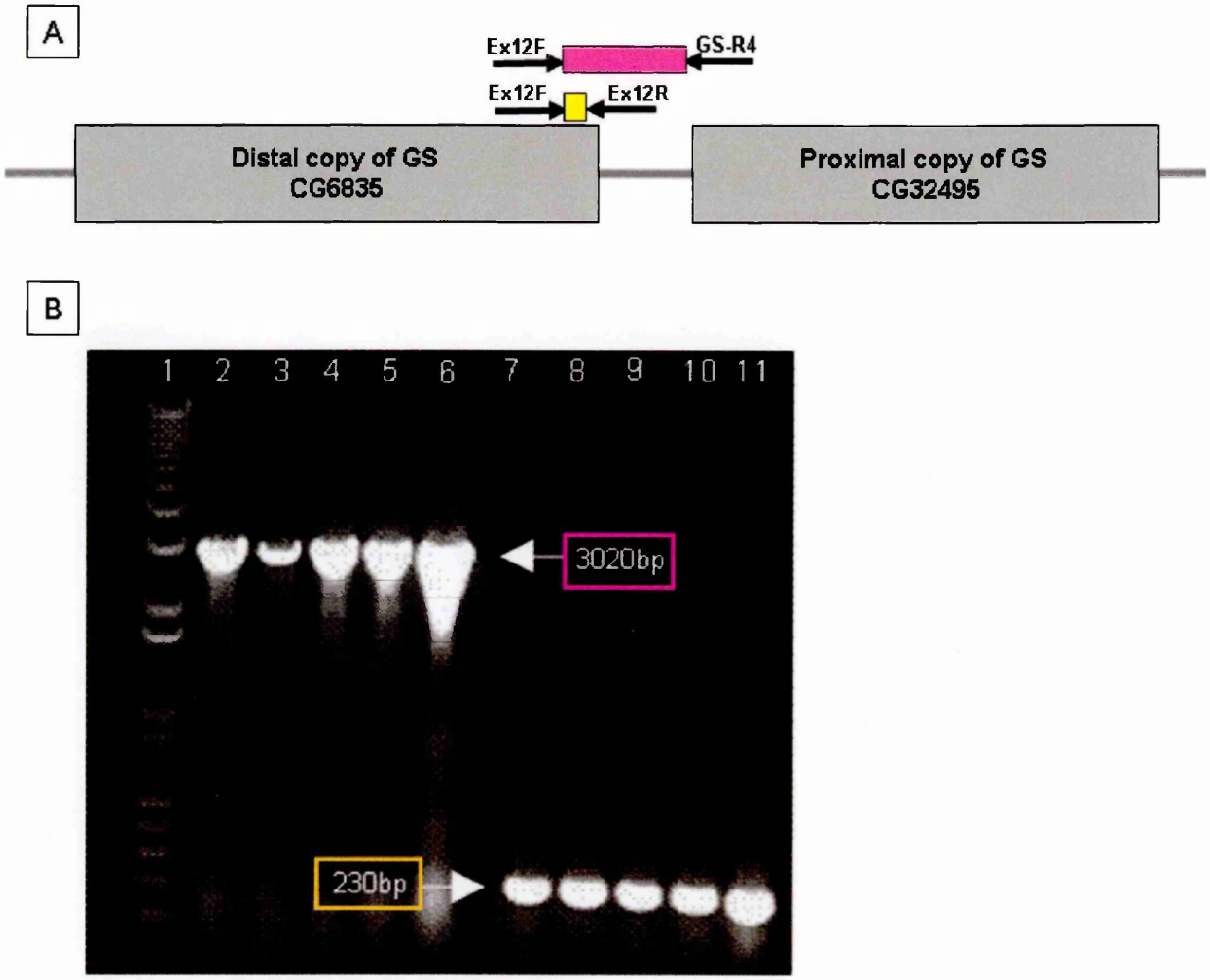


Figure 3.5: The *DmGS* duplication PCR test. Panel A shows a schematic representation of the primers used to amplify the region between the two copies of *DmGS*. Primer Ex12F was designed to anneal to the 3′ end of the distal copy of *DmGS*, and to form a product with the primer GS-R4, which was designed to correspond to the 5′ region of the proximal copy of *DmGS*. The size fragment expected from an amplification of the region between the two copies of *DmGS* was 3020bp. A control PCR reaction was performed to test DNA integrity using primers Ex12F and Ex12R, which amplifies a 230bp region of the 3′ end of *DmGS*, regardless of the duplication status. Panel B is a photograph of an agarose gel with PCR products. DNA was extracted from male flies of each strain tested, and a PCR reaction was performed to amplify the region between the two copies of *DmGS* with primers Ex-12F and *DmGS*-R4. CantonS was used as a positive control for the presence of the duplication in all PCR reactions. PCR cycling conditions used were as described previously, with an annealing temperature of 52°C and an extension time of 2 minutes. Lane 1= 1Kb plus DNA marker, lanes 2-6 contain PCR products from primers 12F-R4 for w^{1118}, CantonS, EP0970, EP1322 and BACR29H04. Lanes 7-11 are the positive controls for the previous six PCR samples amplified with primers Ex12F-12R.

An agarose gel with an example of the PCR reaction using strains with the duplication is shown in Figure 3.5B. It is interesting to note that not all strains of *D. melanogaster* tested contained the *DmGS* duplication (Table 3.1). None of the sibling species tested (*Drosophila mauritiana* C164.1, *Drosophila simulans* Oxnard, *Drosophila simulans*

Ethiopie 225.1, *Drosophila simulans* C167.4, *Drosophila sechellia* Jallons, and *Drosophila sechellia* Cousin) contained the duplication (Table 3.2), so it is suspected the duplication is present only in specific strains of Drosophila melanogaster.

Table 3.1: Results of the PCR test for the presence of the *DmGS* duplication in a range of *D. melanogaster* laboratory strains

D. melanogaster strain	Duplication status
winscy	One copy
l(1)8.3/FM6	One copy
FM6/C(1)DX yf	One copy
FM7	One copy
w sn^{3}	One copy
f	One copy
w m f	One copy
FM6/Df(1)C128	Two copies
y sn^{3} *v*	Two copies
y^{2} *m*^{74F} *wy*^{74i} *sd*^{1} *os*^{s}	Two copies
CantonS	Two copies
w sn^{3} *EP1322*	Two copies
w^{1118} *sn EP0970*	Two copies
S2 cells	Two copies
w^{1118}	Two copies

Table 3.2: Results of the PCR test for the presence of the *GS* duplication in a range of *Drosophila* sibling species

Drosophila species	Duplication status
Drosophila mauritiana C164.1	One copy
Drosophila simulans Oxnard	One copy
Drosophila simulans Ethiopie 225.1	One copy
Drosophila simulans C167.4	One copy
Drosophila sechellia Jallons	One copy
Drosophila sechellia Cousin	One copy

CantonS, a widely used wild type strain, does have two copies of the gene, while a variety of strains held in our laboratory, including the balancers *winscy*, *FM6*, *FM7*, and strains marked with *w sn*^{3}, *f*, and *w m f* only had one copy of *DmGS*. The commonly used *Drosophila* S2 cell line had two copies of *DmGS*. A selection of wild type *Drosophila melanogaster* strains from different geographical regions and collection dates were also tested for the duplication. All of the wild caught strains possess the *DmGS* duplication (Table 3.3).

Table 3.3: Results of the PCR test for the presence of the *DmGS* duplication in a range of *Drosophila melanogaster* wild type strains

Drosophila melanogaster *strain*	Country of collection	Date of collection	Duplication status
VAG1	Athens, Greece	1965	Two copies
BER1	Bermuda	1954	Two copies
RC1	California, USA	1954	Two copies
BS1	Barcelona, Spain	1954	Two copies
PYR2	Pyrenees, Spain	1965	Two copies
Wild 10D	South Carolina, USA	1966	Two copies
CA1	Cape Town, South Africa	1954	Two copies

3.2.3 Resolution of the position of the 297 element in the DmGS gene duplication

A partial *297* element with a 4.26kb internal deletion is assigned to the distal copy of *DmGS* (CG6835) in FlyBase, but because the sequences of the distal and proximal copy are identical, it was unclear how the sequence assembly process had assigned this element specifically to the distal copy. Direct PCR amplification of the *297* element was not used as a strategy because of the high frequency of *297* elements dispersed throughout the genome: any attempt at amplification using a primer pair complementary to a *297* sequence resulted in numerous bands (data not shown).

In order to confirm that the *297* element was indeed inserted in CG6835, BACR29H04 was digested with *Spe*I and *Kpn*I. The *Spe*I site is distal to CG6835, and is unique to the distal copy of *DmGS*. The positions of the restriction enzymes cleavage sites, and the expected sizes of the digestion products are illustrated in Figure 3.6. The positions of the probe and restriction sites, along with expected fragment sizes are listed in Table 3.4. The probe was amplified using the primer pair GS_upstream_forward and GS_upstream_reverse.

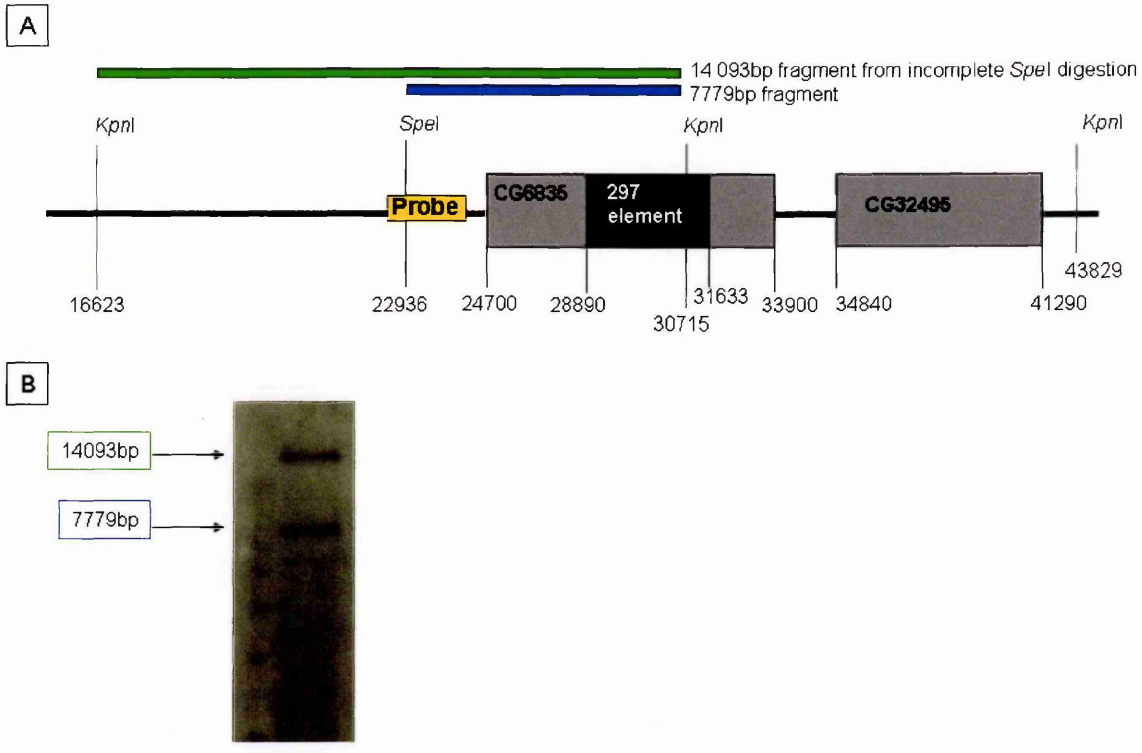


Figure 3.6: Panel A is a cartoon illustrating the positions of restriction enzyme sites relative to *DmGS* and the probe. The sizes of the digestion fragments will depend on whether the 297 element is present in the distal copy of *DmGS* or not. If the 297 element is present in the distal copy (CG6835), the fragment expected will be 7779bp (blue bar). The 14093pb fragment (green bar) is present when there is incomplete digestion of the fragment with SpeI. If the 297 element is absent from the distal copy of *DmGS*, the expected size of the fragment would be 20880bp. The positions of the restriction enzyme sites, as well as the copies of *DmGS* and the 297 element on BACR29H04 (Appendix A) are given on the lower half of the diagram. The position of the probe on BACR29H04 is 22931-23431 (500bp). Panel B shows an X-ray film of the Southern blot performed on BACR29H04, with the DNA fragment expected if the 297 element is in the distal copy of *DmGS*, and there is the large 14093bp fragment produced from an incomplete digestion with SpeI.

Table 3.4: Positions of restriction sites, *DmGS* and the probe on BACR29H04.

Molecular element	Position on BACR29H04
SpeI	22936
KpnI	16623 + 30715 + 43829
Distal copy of *DmGS* (CG6835)	24784-33692
Proximal copy of *DmGS* (CG32495)	34939-41115
Probe	22931-23431
297 element	28890-31633
EP element EP1322	25010

3.2.4 Resolution of the position of P-element EP1322 in the DmGS gene duplication

It was important to unambiguously determine which copy of *DmGS* contained the P-element in strain EP1322. In order to resolve this question, two primer sets were used to amplify the region upstream of the P-element. One set used a forward primer designed

from the unique sequence upstream of the distal copy of *DmGS* (GS-upstream-F), and the second set used a forward primer designed from the sequence at the end of the distal copy of *DmGS* (Ex12F). Both sets used PlacR, which is complementary to the proximal end of the P-element in strain EP1322, as a reverse primer.

The primers GS-upstream-F, which is not in the duplicated region, and Plac-R, were expected to amplify the region upstream of the duplication to the 5′ region of the P-element if the P-element was present in the distal copy (Figure 3.7A). The second primer pair Ex12F and Plac-R would amplify the region from the 3′ end of the first copy of *DmGS* to the 5′ end of the P-element if the P-element was present in the proximal copy. The PCR was performed on DNA extracted from male EP1322 flies, using the CantonS strain as a control. The result from this experiment was an amplification of the 2310bp region between GS-upstream-F and Plac-R (Figure 3.7B), whereas no product was amplified with primers Ex12F and Plac-R (data not shown). These results confirmed that the location of the P-element in EP1322 was in the distal copy of *DmGS*, as is indicated on the sequence data from FlyBase (Celniker et al., 2002).

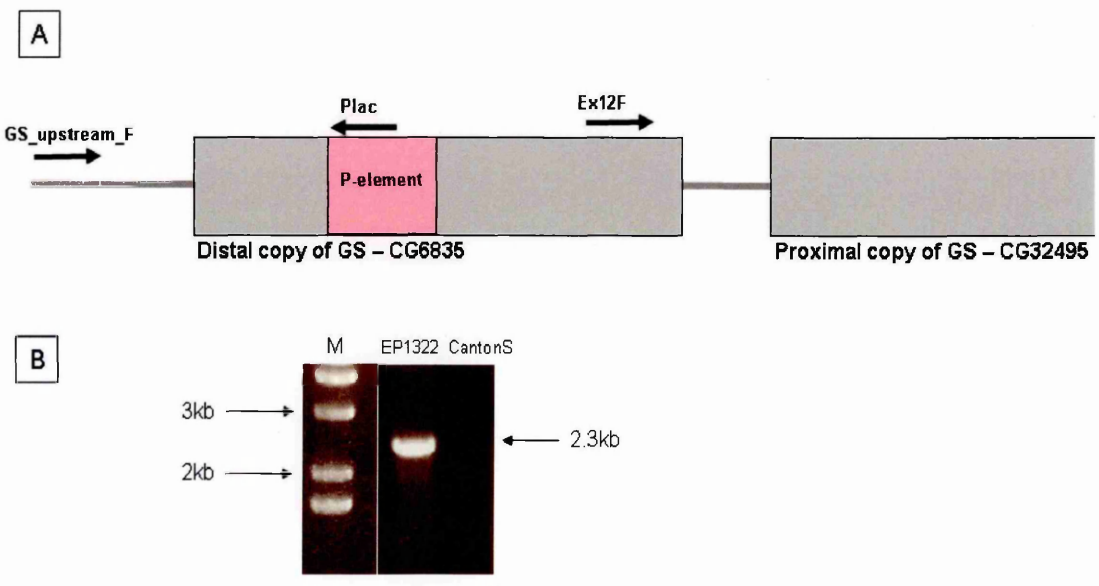


Figure 3.7: Confirmation of the P-element position in the distal copy of *DmGS*. Panel A: Placement of primers on the *DmGS* locus used to determine the position of the p-element EP1322 within the *DmGS* duplication. Panel B: Agarose gel containing the PCR product from primer pair GS_upstream_F and Plac. PCR cycling conditions were standard, with an annealing temperature of 53°C, and an extension time of 3 minutes.

3.3 Discussion

The data in this chapter have clarified the position of the 297 element and the P-element in strain EP1322 . The P-element in strain EP1322, as well as the 297 element, is in the distal copy of *DmGS*, confirming their positions in the published sequence. The transcriptional consequences of the *297* element in an intron of the distal copy of *DmGS* (CG6836) are unclear.

The data presented in this chapter have also demonstrated the presence of a *DmGS* duplication in certain strains of *D. melanogaster.* The majority of *D. melanogaster* laboratory and wild type strains tested for the duplication carried two copies of *DmGS*. The duplication was not, however, present in any of the *Drosophila* sibling species tested. The only *D. melanogaster* strains tested that did not have a duplicated *DmGS* locus specific to *D. melanogaster* were inbred laboratory strains including X-chromosome balancer strains, and the mutant *l(1)8.3/FM6*. The X-linked lethal mutant *l(1)8.3/FM6* was created from the EP1322 parental stock using P-element excision mutagenesis, as discussed further in Chapter 4. The loss of a copy of *DmGS* in *l(1)8.3* could have resulted from the repair of the double strand break induced during P-element excision, by the same mechanism used in targeted mutagenesis to reduce two copies of the target sequence to one (Engels, 1996).

The only other instance of a *GS* duplication in other lineages studied is in the legume family of plants (Frendo et al., 2001). The *GS* locus in legumes contains a tandem duplication of *GS* which occurred through unequal crossing over after the divergence between *Fabales* (legumes), *Solanales* and *Brassicales* families. The legume duplicate *GS* gene encodes two different proteins; one copy encodes a *GS* homologue, and the second copy encodes 'homoglutathione synthetase' which catalyses the formation of homo-glutathione as a result of 2 amino acid mutations (Frendo et al., 2001). Homo-glutathione differs from glutathione in that homo-glutathione contains a β-alanine molecule instead of the glycine residue found in GSH (Matamoros et al., 1999). The two forms of GS have

differing distribution patterns between members of different genera, as well as between different plant organs within one species (Frendo et al., 2001).

Identical tandem gene duplications are thought to arise from unequal cross-over during homologous recombination, where there is a misalignment between genes during meiosis, leading to the formation of gametes with different copy numbers of genes (Lynch and Conery, 2000; Lynch, 2002; Zhang, 2003), illustrated in Figure 3.8.

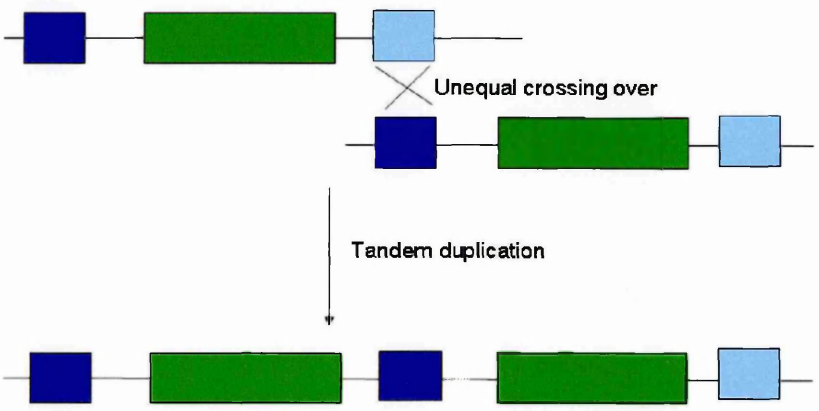


Figure 3.8: Cartoon of the 'unequal crossing over' mechanism of gene duplication. Unequal recombination between paralogous (share a common ancestor due to a duplication event) sequences (light and dark blue boxes) results in a tandem duplication (green box) in the genome of the gamete. Diagram modified from Hurles (2004).

Gene duplications are estimated to arise at a rate one in a hundred genes every million years (Lynch and Conery, 2000), and are thought to play a vital role in the evolution of new gene functions. A large proportion of genes are duplicated in most organisms, with *Drosophila* carrying an average number of duplicated genes, as shown in Table 3.5.

Table 3.5: Gene duplication prevalence in selected organisms. (Zhang, 2003)

Organism	Total number of genes	Percentage of genes arising from duplications
Mycoplasma penumoniae	677	44
Helicobacter pylori	1590	17
Haemophilus influenzae	1709	17
Saccharomyces cerevisiae	6241	30
Caenorhabditis elegans	18 424	49
Drosophila melanogaster	13 601	41
Arabidopsis thaliana	25 498	65
Homo sapiens	30 000	~50

There are four possible fates for a duplicated gene: dose repetition, pseudogenisation, neofunctionalisation or subfunctionalisation. In dose repetition, both copies of the gene remain active. This is beneficial in the case of highly expressed proteins, e.g. histones and ribosomal RNA genes (Hurles, 2004). In pseudogenisation, one of the genes develops a null mutation and becomes a pseudogene (Walsh, 1995; Zhang, 2003), in a process which takes several million years in animals (Lynch, 2002). Neofunctionalisation is the process whereby the duplicated gene either evolves into a new gene with a new function, or, more commonly, a gene with a function related to the parental gene's function (Zhang, 2003). In the *GS* duplication reported in legumes (Frendo et al., 2001), one copy of *GS* retained the original function, and the duplicated copy of *GS* acquired the ability to synthesise homo-GSH.

It has been proposed that it is evolutionarily more favourable for gene duplicates to be conserved by subfunctionalisation (Lynch and Force, 2000), where both copies of the gene accumulate mutations which partition the functions of the single copy gene between the duplicates, necessitating the preservation of both copies (Lynch and Conery, 2000). In this case there is an extended period where both gene copies are exposed to natural selection, enhancing the chance of a beneficial mutation with a new function. One example of subfunctionalisation is the expression of each copy in different tissues and different developmental stages, as is observed in the *Drosophila* β-*tubulin* gene family (Lynch and Force, 2000; Conant and Wagner, 2003). Another interesting case in *Drosophila* is the *Pax*-6 homologue *eyeless*, which is involved in eye development. A duplication of *eyeless* was discovered, and named *twin of eye (toy)*. *Toy* acts upstream of *eyeless* by directly regulating expression of *eyeless* (Czerny et al., 1999). In this case of subfunctionalisation/neofunctionalisation, the duplicated gene has a different function to the parental gene, but is involved in the parental gene network.

It is not known at this stage what effect the presence of two copies of *DmGS* has on strains which carry the duplication, or whether both copies are even functional. It is interesting to note that the sequences of the two copies, according to FlyBase, are identical, including the large stretches of intronic sequence. This would suggest a very recent appearance of the duplication, as the sequence has not diverged at all. Thornton and Long (2002) suggest there was a recent burst of duplication events on the *Drosophila melanogaster* X-chromosomes.

The lack of any difference between the two copies seems to suggest the recent appearance of the second copy of *DmGS*, supported by the lack of the duplication in any sibling *Drosophila* species. It must, however, be noted that the similarity of the sequences in published data could be an artefact of the sequence assembly methods. The presence of a *DmGS* duplication in all tested wild type strains of *D. melanogaster* suggests that the reduction from two copies to one copy of *DmGS* in a few laboratory strains of *D. melanogaster* may have happened recently, perhaps in the last 100 years, in a parental stock kept in captivity. This parental strain could then have been used to create marked stocks that inadvertently carried the single copy *DmGS* mutation.

4 Chapter 4: Characterising a putative *GS* knockout fly

4.1 Introduction

The molecular structure of the glutathione synthetase (*GS*) locus was confirmed in the previous chapter, leading to the next stage in the project: determining the function of *GS* in *Drosophila melanogaster*. A starting technique for determining gene function is a bioinformatic analysis of the region. The locus under investigation in this study has a sequence similarity to GS enzymes from a series of both prokaryotes and eukaryotes, including yeast, worms and mice (Drysdale and Crosby, 2005). After a bioinformatic suggestion of the function of a gene, experimental work involving gene knockouts and overexpression follows.

A powerful method of studying the phenotype of a specific gene is to obtain a mutation in the gene of interest. The effects observed as a result of knocking out the gene of interest can give useful information on the function of the gene. In *Drosophila,* mutants can be created by chemical mutagens or irradiation, but random mutations created using those techniques are not useful when a mutation is required in a specific gene. P-element mutagenesis is a far more valuable technique in the creation of mutations in a particular locus (Adams and Sekelsky, 2002).

P-elements, described in Chapter 1 (section 1.3.2.1), can be used to disrupt gene function by insertion into a gene of interest (Ryder and Russell, 2003). Alternatively, P-elements are used to create mutants by imprecise excision: in approximately 1% of excision events, the break at the excision site will not be repaired correctly, resulting in a deletion of the region surrounding the P-element and a subsequent loss-of-function mutation in a neighbouring gene (Engels, 1996; Ryder and Russell, 2003).

A candidate lethal allele of *DmGS* (*l(1)8.3*) had previously been obtained. This chapter describes work aimed at determining whether *l(1)8.3* was in fact mutant for *DmGS*.

4.2 Results

4.2.1 P-element mutagenesis of DmGS

The attempt to induce a mutation of *DmGS* was started when the existence of a tandem duplication of DmGS was not known (Pushpa Kansagra, pers. comm.). This section reviews the derivation of a candidate lethal allele of *DmGS*. Strain EP1322 contains an EP-element situated in the 5′ end of the distal copy of *DmGS*. An EP element is a modified P-element which carries a UAS near one end (Adams and Sekelsky, 2002). Despite the additional UAS sequence, the EP element functions in the same manner as a non-autonomous P-element would when exposed to transposase, therefore the EP element used in this study will be referred to as a P-element throughout this chapter. The EP1322 P-element was mobilised by a dysgenic cross, using a Δ2-3 transposase element bearing strain (Section 2.30). Occasional mobilisation events would be expected to be imprecise, disrupting a portion of the DmGS gene, and some of these events would be expected to be lethal. In four P- element mobilisation experiments, one X-linked 'rip-out' lethal was recovered (Pushpa Kansagra, pers. comm.).

The crossing scheme used to mobilise the EP1322 P-element is illustrated in Figure 4.1. In the first cross, strain EP1322, marked with *white* (*w*) and *singed* (sn^3) (Pushpa Kansagra, pers. comm.) was crossed to a stock containing Δ2-3 transposase on a chromosome marked with *Stubble* (*Sb*), which enabled mobilization of the P-element. In the second cross, the transposase was eliminated by segregation, and progeny with white eyes were collected as this indicated a loss of the P-element, which was marked with w^+. In the third cross, white-eyed progeny were balanced over *FM6*. The progeny of these single pair matings would be scored for *sn*, which marked the parental EP1322 chromosome, and if males carrying *sn* were absent, it would indicate the presence of a sex-linked lethal mutation. In the case of a sex-linked lethal, female siblings were crossed to *FM6* males to obtain a balanced stock.

The same crossing scheme was performed using strain EP0970 which contains an EP-element approximately 4.5 KB from the 3′ end of the proximal copy of *DmGS*.

Cross (i)

♀♀ w^{1118} *sn*^{3} EP1322 X ♂♂ w^{1118} / Y;*Sb* Δ2-3 /TM6B,*Tb*

↓

Cross (ii)

♂♂ w^{1118} *sn*^{3} EP1322 / Y; *Sb* Δ2-3 /+ X ♀♀ winscy

↓

Cross (iii)

♀ w^{1118} *sn*^{3} EP1322* /winscy; X ♂ FM6

↓single pair matings

If males in the class marked with *sn* indicated below were absent, it indicated a sex-linked lethal

♂ w^{1118} *sn*^{3} EP1322* / Y

If a male lethal class was discovered, the female flies were collected and crossed to FM6 males to obtain a balanced stock

♀ w^{1118} *sn*^{3} EP1322* / FM6

Figure 4.1: The mating scheme followed for imprecise mutagenesis of *DmGS* in stock EP1322. In the first cross strain EP1322 females were crossed to flies carrying a transposase on a chromosome marked with *Sb*. In the second cross male flies carrying a copy of EP1322 as well as the transposase were crossed to winscy X-chromosome balancer to segregate the transposase from EP1322. In the third cross, female flies were crossed to an *FM6* balancer to detect X-linked lethal flies. *EP1322 indicated EP1322 exposed to transposase. The winscy balancer was used in cross (ii) because *FM6/FM6* females are sterile.

Four rounds of imprecise excision mutagenesis using EP1322 were completed, resulting in the recovery of the recessive lethal *l(1)8.3/FM6*. This lethal stock has been the subject of subsequent studies into the gene encoding glutathione synthetase. No lethal mutants were recovered from EP0970 mutagenesis, probably as a result of the larger distance between *DmGS* and the P-element.

4.2.2 Determination of the source of lethality in l(1)8.3/FM6

After the recovery of the lethal mutant *l(1)8.3*, it was important to ascertain whether the lethality was due to disruption of *DmGS* or a result of the disruption of another gene surrounding the P-element after mobilisation of EP1322. The P-element may also have transiently inserted into another region of the genome, where it could interrupt the sequence of an essential gene or its promoter. Three methods are available to determine

whether the lethality in *l(1)8.3* was due to *DmGS* disruption; firstly, the lethal mutation could be mapped by crossing *l(1)8.3/FM6* to a marker stock with a series of easily scorable mutations along the X chromosome. A study of the recombinant classes should then be able to place the source of the lethal mutation close to *l(1)8.3/FM6* or elsewhere along the chromosome. A second approach used was a complementation test using a duplication of a region known to contain *DmGS*. The third and most convincing method to confirm whether the lethality of *l(1)8.3/FM6* is a result of the disruption of *DmGS* would be to rescue the males via a P-element mediated transformation with a transgenic form of *GS*; if the functional transgenic *GS* inserted into the genome allowed males to live, it would prove that *DmGS* was disrupted in *l(1)8.3/FM6*.

4.2.3 Recombination mapping l(1)8.3

In order to carry out this recombination experiment, a stock with markers along the X-chromosome (y^2 m^{74F} wy^{74i} sd^1 os^s) was crossed to *l(1)8.3/FM6*. The crossing scheme illustrated below (Figure 4.2) was followed, and the resultant recombinant flies were scored for the presence of markers. The proportion of recombinant progeny recovered would then indicate where on the X-chromosome the lethal mutation is situated.

♂♂ y^2 m^{74F} wy^{74i} sd^1 os^s	x	♀♀ *w sn* l(1)8.3/FM6
↓ mass mating		
♀♀ l(1)8.3/ y^2 m^{74F} wy^{74i} sd^1 os^s	x	♂♂ FM7/Y
↓		
♀ Recombinant/FM7	x	♂ y^2 m^{74F} wy^{74i} sd^1 os^s /Y
↓ single crosses, score progeny		
Examine progeny: Score recombinant/y^2 m^{74F} wy^{74i} sd^1 os^s ♀ for all markers, and check for viability of recombinant ♂		

Figure 4.2: Recombination mapping of *l(1)8.3*. In the first cross, a multiple balanced stock is crossed to *l(1)8.3/FM6* in mass matings. The *l(1)8.3*/marker females were then crossed to FM7 males. The recombinant marker/*l(1)8.3* was then crossed to the parental marker stock in single pair matings.

This experiment was hindered by the low penetrance of the os^s phenotype, a round, small eye placed more on the top of the head than on the side (Lindsley and Zimm, 1992). The map of the markers in relation to *DmGS* (Figure 4.3) indicates the importance of this marker in determining which parental copy of *DmGS* the recombinant flies carried, as it is the only marker proximal to *DmGS*. Fortunately it was possible to distinguish between the two parental copies of *DmGS* using PCR because the marker stock carried a duplicated form of *DmGS*, while *l(1)8.3* has a single copy of *DmGS*. The number of recombinants recovered in each class is shown in Table 4.1.

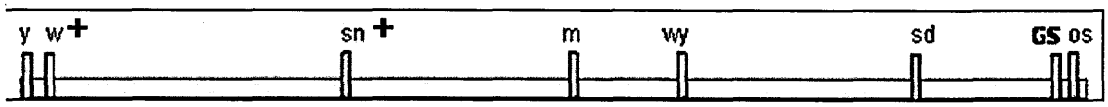


Figure 4.3: Linkage map of the markers present in the marker stock in relation to the markers on *l(1)8.3* and GS. The positions of the alleles on the X-chromosome which has a recombination map length of 66 are: y=0.0, w=1.5, sn=21.0, m=36.1, wy=40.7, sd=51.5, *DmGS*=59, os=59.2

It was unexpected that one of the parental classes (*sn* m^+ wy^+ sd^+) would be so poorly represented, and the reason for this is not known. The majority of the recombinants with the wild type sd^+ phenotype originating from *l(1)8.3* exhibited X-linked lethality, while most *sd* recombinants were not male lethal. The exception is the class 2 recombinant (*sn* m^+ *wy sd*) which was present only as a male lethal stock, where it is possible that a second recombination even occurred between *sd* from the marker stock chromosome, and *DmGS* from the *l(1)8.3* chromosome. *sd* is the closest marker distal to *DmGS*, thus making it less likely for a recombination event to occur between *sd* and *DmGS*. All male flies with the wild type sd^+ allele originating from *l(1)8.3* were then assayed for the *DmGS* duplication, which was present only in the marker stock, using PCR. Every sd^+ recombinant tested had a duplicated *DmGS*, indicating the presence of the wild type *DmGS* originating from the marker stock. The lack of viable males with the mutated form of *DmGS* originating from *l(1)8.3* is good evidence that lethality in *l(1)8.3* is closely linked to the *DmGS* gene.

Table 4.1: Recombinant classes for l(1)8.3/*FM6* crossed to the marker stock y^2 m^{74F} wy^{74i} sd^1 os^s. The total number of vials containing each of the recombinant classes is shown. Of that total, the number of vials containing male lethal flies is listed. All male recombinants with the sd^+ phenotype were tested for the *DmGS* duplication by PCR , and each one of the recombinant males carried the wild type DmGS from the marker stock.

Markers present on recombinant lines					Total number of vials containing the recombinant class	Number of vials containing male lethal recombinants	Number of vials with male flies containing GS duplication
parental	*sn*	m^+	wy^+	sd^+	1	-	1
3	sn^+	m^+	wy^+	*sd*	-	-	-
2	*sn*	m^+	*wy*	*sd*	1	1	-
1	sn^+	m^+	*wy*	sd^+	1	1	-
1	sn^+	m^+	wy^+	sd^+	8	1	7
2	sn^+	*m*	wy^+	sd^+	1	1	-
3	sn^+	*m*	*wy*	sd^+	2	-	2
parental	sn^+	*m*	*wy*	*sd*	30	2	All tested individuals had the duplication

4.2.4 Rescue of l(1)8.3/FM6 by crossing to a strain with DmGS duplication

The mapping of *l(1)8.3* lethality to the *DmGS* locus region in the previous experiment suggested that *l(1)8.3* lethality should be rescued by a duplication including *DmGS*. *l(1)8.3/FM6* was crossed to a strain with a duplication of the region containing *DmGS* on the Y chromosome. The Y chromosome of strain *Dp(1;Y) W73*, y^{31d} B^1B^s / *C(1)DX* y^1f^d/y^1 baz^{EH171} contains a duplication of region 15B1 to 16F of the X-chromosome (Grumbling and Strelets, 2006), which covers region 16F on the X-chromosome containing *DmGS*. *l(1)8.3/Dp(1:Y)W73* males with white eye and singed hair phenotypes would indicate a rescue of *DmGS* however once the experiment was performed, it was found that *l(1)8.3* was not rescued by the duplication *Dp(1:Y)W73* (Table 4.2).

Table 4.2: Crossing scheme to attempt rescue of *l(1)8.3/FM6* by a stock containing a duplication of *DmGS* on the Y chromosome (*Dp(1;Y) W73, y^{31d} B^1B^s / C(1)Dx y^1f^1/y^1 baz^{EH171}*). The genotypes of male and female gametes are shown to illustrate the four possible progeny outcomes. The genotype in red is that of the male fly that would be present in the case of a successful rescue by the duplicated chromosome region. Over 200 individual progeny were scored in this experiment.

		Female gametes	
		FM6	***l(1)8.3***
Male gametes	***y baz***	*y baz / FM6*	*l(1)8.3 / y baz*
	Dp(1;Y)W73	*FM6 / Dp(1;Y)W73*	***l(1)8.3 / Dp(1;Y)W73***

The lack of rescue could be a result of inaccurate mapping of the duplicated portion of the X-chromosome in strain *Dp(1;Y) W73, y^{31d} B^1B^s / C(1)Dx y^1f^1/y^1 baz^{EH171}*, or it could mean that the lethality of *l(1)8.3* is not in fact a result of *DmGS* disruption.

4.2.5 Attempted rescue of l(1)8.3 by complementation with a cloned copy of human GS

The failure to complement *l(1)8.3* with a duplication expected to contain a copy of the *DmGS* locus led to a third strategy to resolve the issue of whether the lethal was mutant for GS. In order to convincingly prove the linkage of lethality in *l(1)8.3* to *DmGS*, a cloned functional copy of a *GS* gene would have to be introduced into a *l(1)8.3* fly, and be shown to reverse the X-linked lethality. In this experiment, human *GS* was used to create a transgenic fly because previous attempts to clone genomic *DmGS* had been unsuccessful due to the large size of the gene and the complexity of the genomic arrangement. At the time of creating the transgenic fly, there was no information available on which of the *DmGS* cDNA clones were functional, or which of the duplicated copies of *DmGS* was functional.

Human GS (*hGS*) cDNA in the bluescript pSK vector was provided by Lesley McLellan (Dundee University). The 1.8kb *hGS* cDNA (Appendix C) was excised from the

Bluescript vector using the restriction enzymes *Xba*I and *Xho*I, and was cloned in the multiple cloning site of the pUAST vector, behind the UAS element (Figure 4.4)

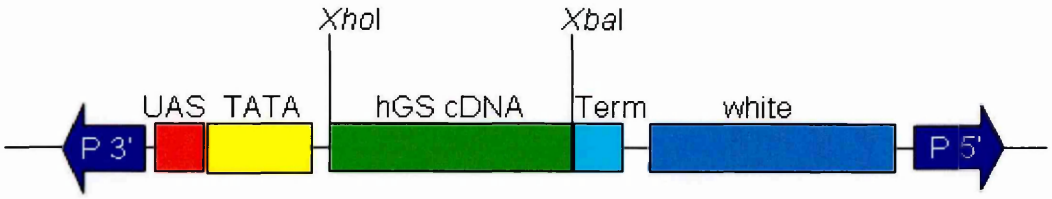


Figure 4.4: A schematic diagram of *hGS* cloned into the pUAST vector. hGS-pUAST was marked with a *white* gene, allowing easy scoring of transformants. Transcription of *hGS* would be driven by the UAS promoter on activation by a GAL4 molecule. The length of the *hGS* cDNA was 1.8kb, making the hGS-pUAST plasmid a total of 10900bp.

Positive clones containing the hGS-pUAST construct were verified by PCR amplification of an insert of the correct size using a primer pair corresponding to the 5′ and 3′ end of the pUAST multiple cloning site (pUAST_F_seq and pUAST_R_seq). The orientation of the hGS insert in pUAST was established by sequencing using the pUAST_F_seq and pUAST_R_seq primer pair.

The transgenic fly line was created by embryo microinjection, as described in Section 2.27. One viable transgenic line was established, and further unique insertions of the *hGS* transgene were obtained by remobilisation in order to recover a greater number of insertion points, as a means to eliminate the 'position effect' (Tower, 2002). The crossing scheme below describes the P-element remobilisation (Figure 4.5). Three separate lines were created from this remobilisation event (lines A, B and C).

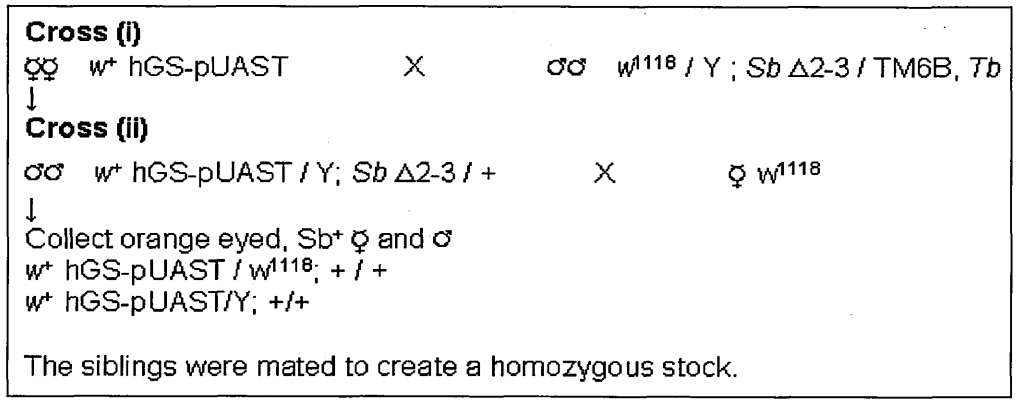


Figure 4.5: Mapping the transgene hGS-pUAST. In the first cross, homozygous hGS-pUAST females were crossed to male flies carrying a transposase on a chromosome marked with *Sb*. In the second cross, males carrying hGS-pUAST as well as the transposase were crossed to w^{1118} virgins, enabling the segregation of the transgene and the transposase in the following generation. Flies which lacked transposase (Sb^+) which carried hGS-pUAST (w^+) were collected, and crossed to create a remobilised homozygous stock. A darker orange eye colour indicated a double copy of the transgene, and therefore homozygotes.

Following remobilisation, the three hGS-pUAST insertions were mapped to chromosomes by segregation analysis (Figure 4.6). Briefly, the transgenic flies were crossed to a stock marked on the second and third chromosomes, allowing the linking of the transgene (marked with w^+) and a marker specific to a particular chromosome. The insertion of a transgene onto the fourth chromosome is very rare because of the small size of this chromosome, so it was not tested in this study. Of the three separate insertion events created with the remobilisation, two were on the X-chromosome (lines B and C), and one (Line A3) had the transgene inserted on the second chromosome. The hGS insertion in a second line (A4) originating from the original insertion point 'A' also mapped to the second chromosome.

Cross (i)

♀♀ w^+ hGS-PUAST X ♂♂ w^{1118} / Y ; CyO/Bl; TM6*Tb*/TM2*Ubx*

↓

Collect ♂ with red eyes (hGS-pUAST)

Cross (ii)

w^+/Y; Cyo/?; TM6*Tb*/?

w^+/Y; Bl/?; TM2*Ubx*/? X ♀ w^{1118}

If hGS-pUAST was on the:

X chromosome: there would be no red eyed ♂ progeny

2nd chromosome: there would be no red eyed flies marked with Cyo or Bl

3rd chromosome: there would be no red eyed flies marked with *Tb* or *Ubx*

Figure 4.6: Crossing scheme to determine which chromosome the transgene hGS-pUAST was inserted into. In the first cross, homozygous hGS-pUAST females were crossed to male flies carrying a dominant marker on each chromosome. Male progeny which carried the transgene (red eyes) in an unknown chromosomal position, shown by '?', were crossed to w^{1118} females. The resulting progeny would give information on the chromosomal positioning of the transgene.

In order to drive expression of hGS-pUAST, it was necessary to introduce an appropriate GAL4 driver element. GAL4 would bind to the yeast UAS promoter upstream of hGS, thereby driving hGS expression in the tissues where GAL4 was expressed (Duffy, 2002). A stock with GAL4 fused to a heat shock promoter was used as the source of GAL4 to cross into the hGS-pUAST transgenic line. This stock, w^*;*P{*w^{+mC}*=GAL4-Hsp70.PB}2/CyO*, was previously shown to have a 'leaky' heat-shock promoter, and GAL4 expression was shown even in the absence of a heat shock being administered to the flies (Ruth Akhtar, pers. comm.). Even though the heat shock promoter was known to be active in the absence of a heat shock, the crosses using this fly stock were exposed to a heat shock to ensure high activity. The crossing scheme to obtain a *l(1)8.3/FM6*; GAL4-HS/GAL4-HS fly is illustrated in Figure 4.7 overleaf.

To make *l*(1)8.3/FM6;Gal4/Gal4, the progeny of cross(ii) was mated with the progeny of cross(iii) in cross(iv).

Cross (i)
♂ *w*/Y;CyO/*Bl* x ♀ *l*(1)8.3/FM6
↓
Cross (ii)
♂ FM6/Y;*Bl*/+ [from (i)] x ♀ *l*(1)8.3/FM6
↓
♀ FM6/*l*(1)8.3;*Bl*/+ x ♂ *w**/Y; P{*w*+mc=Gal4-Hsp70PB}/Cyo

Cross (iii)
♂ FM6/Y; CyO/+ [from (i)] x ♀ *l*(1)8.3/FM6
↓
Cross (iv)
♀ *l*(1)8.3/FM6; CyO/+ [from (iii)] x ♂ FM6/Y; *Bl*/P{w+mc=Gal4-Hsp70PB} [from (ii)]
↓
Collect male and female progeny, marked with Cyo, and *Bl*/+ and cross to make a balanced stock.

♀ *l*(1)8.3/FM6; CyO/Gal4
♂ FM6/Y; CyO/Gal4

In the next generation, flies homozygous for Gal4 were collected, resulting in the stock:
***l*(1)8.3/FM6; Gal4/Gal4**

Figure 4.7: Crossing scheme to introduce a GAL4-HS driver into *l(1)8.3/FM6*. First, the second chromosome of *l(1)8.3/FM6* was marked with Bl or CyO by crossing to a Bl/CyO stock. The progeny of cross (i) was used in cross (ii) as well as in cross (iv). In cross (ii) the GAL4-HS strain was crossed to *l(1)8.3*/FM6*FM6* marked with Bl. The progeny from that cross was mated to the progeny from cross (iii) in cross (iv). The male and female progeny from cross (iv) which were marked with CyO were collected and crossed to create a stock homozygous for the heat shock GAL4 driver on the second chromosome.

Expression of GAL4 in *l(1)8.3/FM6*; GAL4-HS/+ flies

To test whether the GAL4-HS driver in the *l(1)8.3/FM6*; GAL4-HS stock was capable of driving a UAS-transgene, flies were crossed to two different stocks containing GFP-UAS transgenes. The first stock (w^*; P{w^{+mc}=UAS-GFP.S65T}T2) ubiquitously expressed GFP when driven by GAL4, and the second stock (w^{1118}; P{w^{+mc}UAS-GFP.NLS}14) expresses a GFP with a nuclear localisation signal. By following the pattern of GFP expression once crossed to *l(1)8.3/FM6; GAL4-HS* , the pattern of where the transgenic *hGS* expression could be inferred.

F_1 larvae were observed under a fluorescent microscope at first, second and third larval instar stage. It was not possible to view adult flies under the fluorescent microscope because the adult cuticle did not allow penetration of the GFP fluorescent signal, and there was auto-fluorescence from the cuticle itself. The results show that the GAL4-HS driver was expressing ubiquitously (Figure 4.8). The parental UAS-GFP stocks did not fluoresce (Figure 4.8A and B), but when strain w^*;P{w^{+mc}=*UAS-GFP.565T}T2* was crossed to *l(1)8.3/FM6*;GAL4-HS, there was a high level of expression throughout the larval body (Figure 4.8C and D). The UAS-GFP-NLS stock crossed to *l(1)8.3/FM6*;GAL4-HS had an unusual pattern of expression, seemingly localised to the salivary glands, and trachea (Figure 4.8E). This could be an effect of the NLS tag on the GFP. This experiment confirms the efficacy of the GAL4 driver chosen to drive expression of the hGS-pUAST transgene.

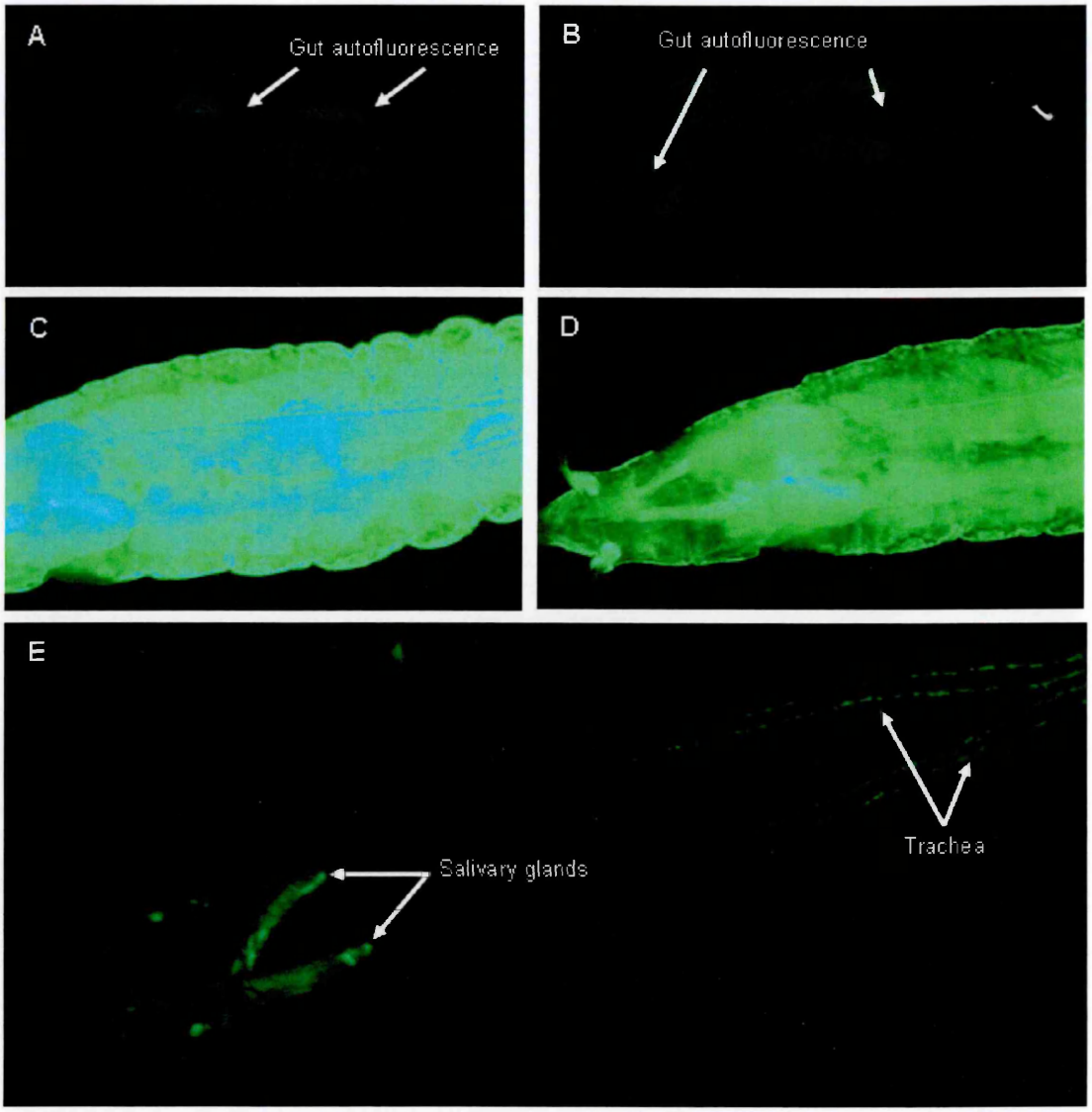


Figure 4.8: Fluorescent microscopy of second instar larvae from *l(1)8.3/FM6*; GAL4-HS crossed to two different GFP-UAS sources. Panel A is the parental strain w^{*};P{w^{+mc}=UAS-GFP.S65T}T2, and Panel B is the parental strain w^{1118}; P{w^{+mc}UAS-GFP.NLS}14. In Panel A and B, the exposure was for 2 seconds. Panels C and D are second instar larvae from *l(1)8.3/FM6*;GAL4-HS crossed to w^{*};P{w^{+mc}=UAS-GFP.S65T}T2. Panel C is a second instar larva with a 2 second exposure, and Panel D is the same larva, but with a 100millisecond exposure. Panel E is a third instar larva from *l(1)8.3/FM6*;GAL4-HS crossed to w^{1118}; P{w^{+mc}UAS-GFP.NLS14}, with a 500 millisecond exposure. Embryos and larvae were heat shocked at 30°C for 30 minutes each day.

In order to ascertain whether expression of a human GS transgene could rescue *l(1)8.3*, *w*/Y;hGS-pUAST flies were crossed with *l(1)8.3/FM6*;HS-GAL4/*CyO*. Second chromosome hGS-pUAST insertion lines A3 and A4 were used for the rescue experiment illustrated in Figure 4.9 below. The presence of B^{+} males of the genotype *l(1)8.3*/Y;HS-GAL4/hGS-pUAST would indicate rescue of *l(1)8.3* by hGS. There were, however, no male flies recovered with this genotype in over 200 individual progeny scored. This result cast further doubts on whether *l(1)8.3/FM6* is indeed a *DmGS* mutant, but because the

failed rescue may have been due to incorrect expression of hGS, further analysis of the hGS transgenic fly was required.

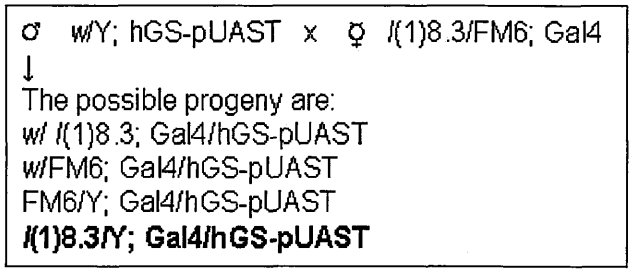


Figure 4.9: Cross to test whether a transgenic copy of humanGS could rescue the putative *DmGS* lethal *l(1)8.3*. Males from a homozygous stock of *hGS*-pUAST were crossed to female *l(1)8.3/FM6* flies. The presence of a male progeny without Bar eyes (marker on *FM6*) would indicate a successful rescue of the lethal. The parental stocks and embryos were heat shocked at 30°C for 30 minutes each day.

The apparent failure of *l(1)8.3* to be rescued by expression of an hGS transgene could be explained in three ways. Firstly, the transgene might not express hGS, secondly the expression pattern of the hGS transgene might be inappropriate for rescue of a *DmGS* mutant, and finally, the lethality of *l(1)8.3* may be unconnected with defects in *DmGS*. To clarify the situation, western blotting analysis of transgenic flies was undertaken. Protein was extracted from *l(1)8.3/w*; HS-GAL4/hGS-pUAST flies. The hGS-pUAST stock was also crossed to a variety of GAL4 drivers, and protein extracted for western blot analysis (Table 3). No bands were present on the blot, indicating that no hGS protein was being expressed in the driven hGS-pUAST flies, using the HS-GAL4 driver or any of the other drivers.

Table 4.3: List of GAL4 drivers crossed to hGS-pUAST to test protein expression of hGS from the transgene.

Driver element	Expression pattern of GAL4
w^{*}; $P\{w^{+mC}$=*GAL4-Hsp70.PB*$\}2$/*CyO*	Heat shock GAL4 - Ubiquitous
y^{1} w^{*}; $P\{w^{+mC}$=*tubP-GAL4*$\}$*LL7/TM3*, Sb^{1}	Tubulin
$y^{1}w^{*}$; $P\{w^{+mC}$=*Act5C-GAL4*$\}$*25FO1/CyO*	Actin
w^{-}; ; $P\{w^{+}$, *GAL4*$\}$-*D42*	Motorneuron

The lack of protein expression from the hGS-pUAST transgene was an unexpected result, and it is still not clear why no protein was expressed. The most likely reason is that the cDNA sequence was cut directly from the bluescript plasmid and cloned into pUAST, without the addition of a translation initiation sequence. The translation initiation sequence

specific to *Drosophila* is known as a Cavener sequence (Cavener, 1987). The sequence which should precede the ATG start codon in *Drosophila* to initiate translation is **(A/C)AA(C/A)**ATG. The sequence preceding the start codon in the human GS cDNA used to create the hGS-pUAST transgene was **TGGG**ATG. The human cDNA therefore did not contain the Cavener sequence, and it may be the cause of the lack of translation of the transgenic protein. For future cloning attempts, the human cDNA should be amplified using primers which would incorporate the Cavener sequence into the *hGS* sequence (Ullmann et al., 1996).

4.2.6 Attempted rescue of l(1)8.3/FM6 by glutamate-cysteine-ligase (GCL)

After the unsuccessful rescue of *l(1)8.3* by hGS, rescue of the putative *DmGS* mutation was attempted with glutamate cysteine ligase (GCL). GCL holoenzyme is the precursor to GS in the glutathione synthesis pathway, and is the rate-limiting step in the reaction. GCL is a heterodimer which consists of a catalytic subunit (GCLC) and a modifier subunit (GCLM). GCLC is functional in isolation, however its activity is enhanced in the presence of GCLM. In *S. cerevisiae*, incubation with the intermediate product produced by GCL, γ-glutamylcysteine, can rescue a glutathione deficiency caused by a mutation in either GCL or GS, thereby negating the need for GS (Grant et al., 1997). If this is true in *Drosophila*, overexpresing GCL might be expected to rescue a *DmGS* mutant.

l(1)8.3/FM6;GAL4-HS was crossed to two transgenic flies; the first overexpressed GCLC only, and the second over-expressed both GCLC and GCLM. The GCLC-pUAST and GCLC/GCLM-pUAST transgenic flies were both created by Pushpa Kansagra (pers. comm.). The crossing scheme for the attempted rescue of *l(1)8.3* by overexpressing GCL is illustrated in Figure 4.10 below:

♂ w/Y; GCLC-pUAST
Or w/Y; GCLC-pUAST; GCLM-pUAST x ♀ *l*(1)8.3/FM6; Gal4
↓
The possible progeny are:
♀ *l*(1)8.3/*w*; Gal4/GCL
♀ FM6/*w*; Gal4/GCL
♂ FM6/Y; Gal4/GCL
♂ ***l*(1)8.3/Y; Gal4/GCL**

Figure 4.10: Crossing scheme for attempted rescue of *l(1)8.3* by overexpressing GCL. Males homozygous for GCLC-pUAST or homozygous for GCLM-pUAST;GCLC-pUAST were crossed to female *l(1)8.3/FM6*;GAL4-HS flies. The presence of a male progeny without Bar eyes (marker on *FM6*) would indicate a successful rescue of the lethal. The parental stocks and embryos were heat shocked at 30°C for 30 minutes each day.

The presence of B^+ eyed males with the genotype *l(1)8.3/FM6*; GAL4/UAS-GCL would indicate rescue of *l(1)8.3* by GCL. There were, however, no male flies recovered with this genotype. This outcome is ambiguous, however, because the lack of rescue could be a result of γ-glutamylcysteine not being sufficient to cover for decreased GSH levels in a *Drosophila GS* mutant, or it could be because *l(1)8.3* is not a *DmGS* mutant.

4.3 Discussion

This chapter reports work aimed at establishing whether or not a recessive lethal derived from the mobilisation of a P-element inserted in the distal copy of *DmGS* indeed resulted from loss of GS expression. It is known that when a P-element is mobilised, the transposition creates a double strand break (Engels, 1996). The majority of the repair events following the mobilisation involve homologous recombination with a sister chromatid, resulting in a replacement of the excised P-element, while approximately 15% of the repair events use a homologous chromosome (Engels, 1996). Imprecise excision of a P-element is a consequence of aberrant repair, and often results in a deletion of the P-element and a portion of the sequence surrounding the element (Figure 4.11). Because P-elements often insert between the promoter and the protein coding region of a gene, imprecise excision can lead to a complete loss-of-function mutation in the gene containing the P-element (Adams and Sekelsky, 2002). The frequency of imprecise excisions is approximately 1% of the excision events (Ryder and Russell, 2003).

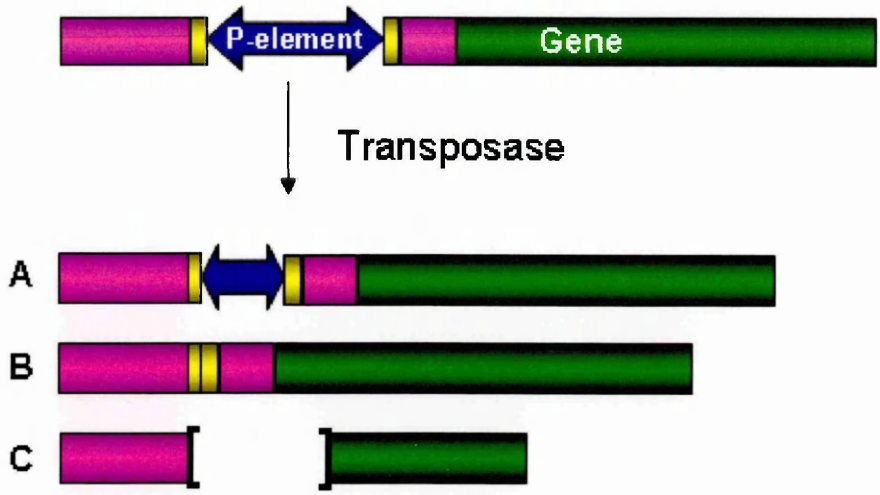


Figure 4.11: Three possible outcomes from P-element mutagenesis. The parental P-element construct is shown at the top of the diagram. Pink bars indicate the promoter/intergenic region, yellow bars indicate the 8-bp target site duplication, green bars indicate the coding region of the gene, and the blue arrow is the P-element with inverted repeats at the ends. 'A' illustrates a homologous recombination event from a sister chromatid which results in a reproduced P-element, occasionally with an internal deletion. 'B' illustrates a homologous recombination event from a homologous chromosome where the P-element undergoes 'precise excision'. 'C' is an example of 'imprecise excision' which arises from a non-homologous end joining event after the ends of the break have been degraded, resulting in a deleted portion of the gene of interest. Adapted from Adams and Sekelsky (2002).

The hypothesis of the mobilisation experiment was that since EP1322 is situated in the 5'end of *DmGS*, its transposition resulted in an imprecise excision in *l(1)8.3*, leading to a null mutation in *DmGS*. The deletion of a portion of the *DmGS* locus surrounding EP1322 in *l(1)8.3* is supported by the reduction in *DmGS* copy number from two to one in *l(1)8.3*, as described in Table 3.1. The strain in which *l(1)8.3* was induced (EP1322) has two copies of *DmGS*, suggesting that the P-element transposition caused a deletion of the distal copy of *DmGS* in *l(1)8.3*.

One of the possible mechanisms whereby the two copies of *DmGS* could be reduced to one copy in *l(1)8.3* is homologous recombination. A possible scenario is illustrated in Figure 4.12, where a double strand break induced by the transposed P-element is repaired by homologous recombination using the duplicated copy of *DmGS* on the same chromatid as the break. The effect of this is the reduction of the tandem duplication to one copy. This mechanism of reducing a tandem duplication to one copy by homologous recombination following a double strand break has been documented previously in a gene targeting system (Rong and Golic, 2000) .

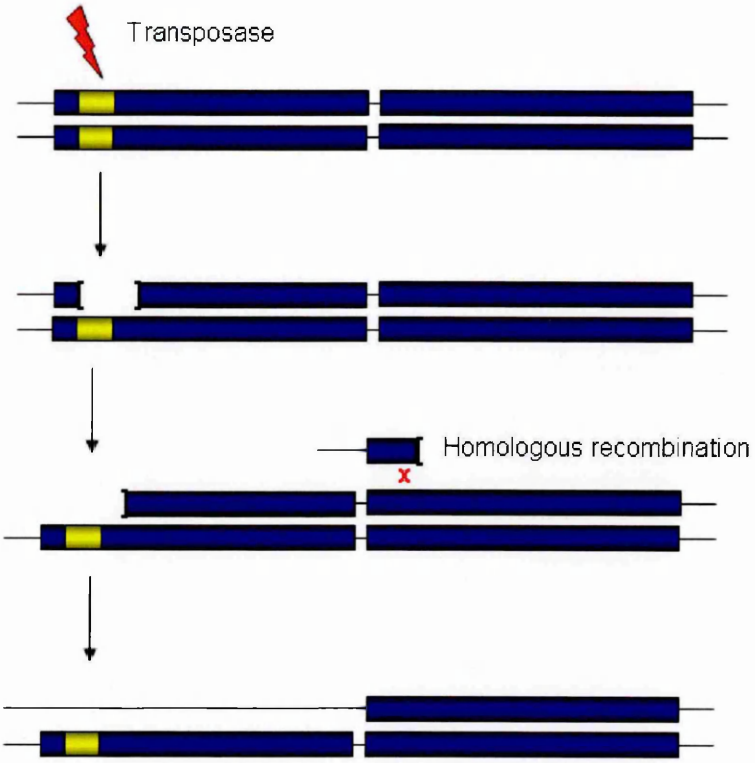


Figure 4.12: A possible mechanism for the reduction of GS copy number from two to one in *l(1)8.3* following P-element mutagenesis. The parental EP1322 strain sister chromatids each had a duplicated copy of GS (blue) and a P-element situated in the 5′end of GS (yellow). After exposure to transposase, the P-element was excised, leaving a double strand break. The break was then repaired by homologous recombination using the duplicated copy of GS on the same chromatid as the excised P-element, rather than the sister chromatid. The net effect of this recombination event was the loss of the first copy of GS, creating a single copy of the gene where there was a duplication.

The most robust way to test the hypothesis that *l(1)8.3* was a *DmGS* mutant would be by complementation with a GS transgene. Previous studies have expressed human SOD (Parkes et al., 1998) and catalase (Ha et al., 2005) enzymes in *Drosophila* with success, achieving lifespan extension as well as elevated oxidative stress resistance in the transgenic flies. Human GS and *Drosophila* GS proteins share 35% identity over the whole sequence, with three out of the four active site residues conserved (Marchler-Bauer and Bryant, 2004). Because of the high homology in conserved regions between *Drosophila* and human GS proteins, it was thought that the human GS protein should have functional activity when expressed in *Drosophila* (Figure 4.13).

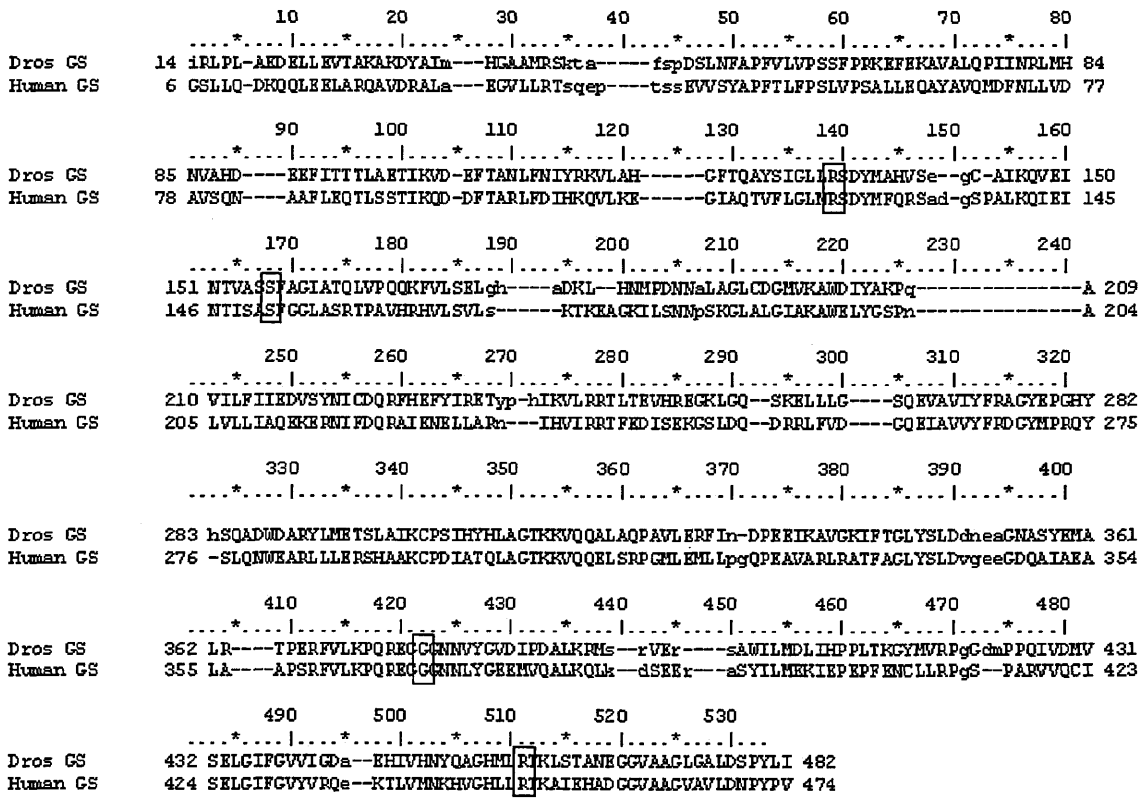


Figure 4.13: ClustalW alignment of human GS and *Drosophila* GS (CG 32495B) protein. The active site residues are marked with a box. The alignment was performed using BioEdit (Hall, 1999).

The GAL4-UAS system used to drive expression of a *hGS* transgene is described in detail in Chapter 1. The *hGS* cDNA cloned into a pUAST vector should have led to overexpression of *hGS* when crossed to a fly expressing GAL4. Unfortunately no expression of hGS was detectable, using a UAS-hGS element in combination with four different GAL4 driver stocks, each expressing GAL4 in different tissues. The lack of *hGS* expression confounded interpretation of the complementation experiment. One of the possible causes of no protein expression could have been that the *hGS* cDNA was cloned into pUAST directly from a vector, without the addition of a translation initiation sequence upstream of the start codon by PCR amplification before cloning. Although this has not been a problem in our laboratory before (Pushpa Kansagra, pers. comm.) it could be the case that for this particular cDNA, an initiation sequence was essential for the protein to be produced.

The question of whether *l(1)8.3* is a *DmGS* mutant is still unresolved. Evidence linking *DmGS* to the lethality in *l(1)8.3* came from the mapping experiment, which suggested that the lethality in *l(1)8.3* was linked to *DmGS*. Also supporting the linkage of *DmGS* and *l(1)8.3* lethality was the observation of GS copy number reduction in *l(1)8.3*. The hypothesis that *l(1)8.3* lethality is linked to *DmGS* was countered by the lack of rescue by a stock carrying a duplicated region of *DmGS* (16F1 on the X-chromosome) on the Y-chromosome. The stock with the duplicated *DmGS* region listed on FlyBase as having a duplication covering the region of 15BD to 16F. There is, however, a note attached to this stock on the Bloomington Stock Centre record raising the possibility of non-continuous coverage around the region of 15BD to 16F on the X-chromosome in the duplicated stock (Eberl et al., 1992). The *DmGS* duplication is situated in the 16F1 region of the X-chromosome, from position 17723484-17732392 (distal copy) to 17733676-17739814 (proximal copy) on the annotated sequence. It is possible therefore that the region intended to cover *DmGS* in the duplicated stock is missing the portion deleted from *DmGS* in *l(1)8.3*. This would explain the unsuccessful rescue seen in this experiment.

The failed rescue of *l(1)8.3* by GCL overexpression was also inconclusive, and may have been because the product of GCL overexpression, γ-glutamylcysteine, was not functionally capable of rescuing a glutathione deficiency in *Drosophila*, as is the case in *S. cerevisiae* (Grant et al., 1997). The ability of γ-glutamylcysteine to act as an antioxidant in the place of glutathione has only been observed in *S. cerevisiae*, making it conceivable that *Drosophila*, as with other multicellular organisms, requires glutathione in addition to γ-glutamylcysteine.

On balance, although a lethal mutation in *DmGS* in strain *l(1)8.3* can not be ruled out, it is likely that the lethal mutation is not in *DmGS*, but is rather a mutant of a gene in the region surrounding *DmGS*. The region neighbouring *DmGS* in 16F on the X-chromosome is a well-documented haplolethal region where a high percentage of the genes have vital functions compared to the surrounding loci outside the gene cluster (Prado et al., 1999).

The region, known as the *Shaker* region is approximately 350kb long, and consists of a dense cluster of genes with a close functional relationship, where a deletion in one gene would inactivate the whole set of related genes. It is possible that the imprecise mutagenesis of EP1322 was associated with mutation of a locus in the haplolethal region. It has been established by PCR that a portion of the distal copy of *DmGS* was removed in *l(1)8.3*, however it may not be the cause of the lethality in *l(1)8.3*. Further evidence for this will be discussed in Chapter 6, where glutathione levels in this strain balanced over *FM6* were found to be equivalent to other fly strains tested, instead of the expected lower glutathione levels in the case of an inactive copy of *DmGS*.

5 Chapter 5: Analysis of predicted *DmGS* transcripts

5.1 Introduction

Release 4 of the *Drosophila* genome sequence assigns six transcripts to *DmGS*: three each are assigned to CG6835 and CG32495. In addition to these six transcripts, another transcript present in an earlier release of the sequence has subsequently been removed. Our collaborators at Dundee University (Shanks, 2004) have isolated a further cDNA transcript, and one further transcript discussed in this chapter was isolated in this study.

The presence of these nine possible transcripts raised two questions. Firstly whether all these transcripts are present *in vivo*, and secondly, what the biological functions of all the transcripts are. The existence of multiple *GS* transcripts has previously only been observed in the mouse (Shi *et al.*, 1996). Murine GS is encoded by a single gene, and has six different transcripts, three of which encode identical proteins and differ only in their transcription start sites (A1, A2, A3). Two of the transcripts encode a second protein, but differ in their transcription start sites (B1, B2), and the final mouse *GS* transcript (C) is unique (Shi *et al.*, 1996). It is interesting to note that of the three transcript classes, only class A produced a protein which could complement a GS deficient *S. pombe* strain, leaving unanswered the question of the function of the remaining two classes of transcripts which did not produce an active GS enzyme, although they were present in specific tissues. This is one of the questions that will require further study in mice as well as in *Drosophila.*

In this chapter, the nine proposed transcripts for *Drosophila* GS are studied to try to determine which transcripts produce an active protein. Several strategies are used: a bioinformatic approach, a molecular approach by cloning and sequencing of *DmGS* cDNAs, as well as a genetic complementation study in a GS deficient strain of *S. pombe* (Mutoh and Hayashi,

1988). Finally, the temporal and spatial pattern of *DmGS* transcription in *Drosophila* was studied.

5.2 Results

5.2.1 Predicted Drosophila GS transcripts

During the course of this project, the sequence of the *DmGS* locus was released for the first time, with transcripts attached to each of the *DmGS* gene copies. The nine transcripts so far predicted are illustrated in Figure 5.1. T1, T2 and T3 are the transcripts which were assigned to the distal copy of *DmGS*, CG6835, and T6, T8, and T9 are assigned to the proximal copy of *DmGS* called CG32495 in Release 4.3 of the *Drosophila* Genome Sequence (Grumbling and Strelets, 2006). Transcript T7 was originally assigned to CG32495 in the Release 3.0 of the *Drosophila* sequence, however this particular transcript has been removed from subsequent annotated versions. Transcript T4 was isolated by Shanks (2004), and transcript T10 was isolated in this study as a new transcript.

The size of the predicted proteins from each of the full length transcripts is in the range of 491-650 amino acids, as listed in Table 5.1. The two truncated proteins from T3 and T9 are predicted to be 127aa, and 308aa respectively. The proteins encoded by each transcript will be discussed in more detail in section 5.2.4.

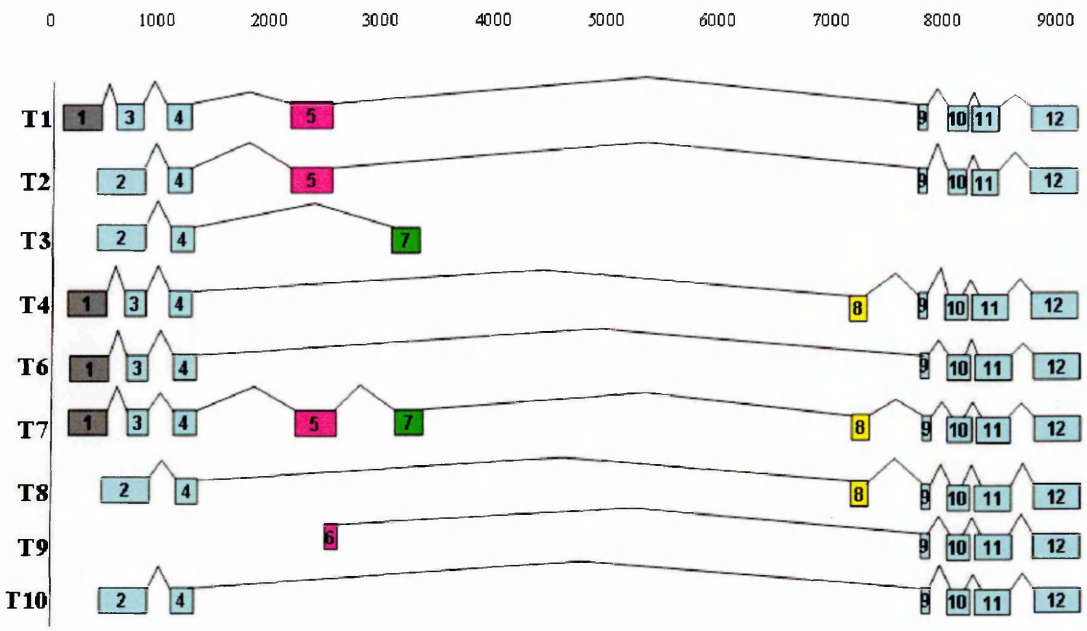


Figure 5.1: A diagrammatic representation of the transcripts predicted for *DmGS* as well as the neighbouring CG32495. Specific exons are indicated by numbers 1-12. The genomic region length is shown by the scale at the top of the diagram (0bp to 9000bp). Transcript 4 was isolated as a cDNA clone by our collaborators in Dundee (Shanks, 2004). A second cDNA transcript (T5) recovered by this group was identical to T2, and has therefore been removed from this schematic. Transcripts T1 and T2 correspond to transcripts CG6835 RC and RD respectively, and T3 corresponds to CG6835 RA. T6, T8 and T9 correspond to CG32495 RA, RB and RC. Transcript 7 was removed from FlyBase in release 3 of the annotated sequence. There is EST coverage of all of the exons, however no complete cDNA sequences had been submitted for any of the transcripts before this study.

Table 5.1: Sizes of the predicted proteins from all nine transcripts.

Transcript/ Class	Transcript name on Flybase	Predicted protein size (aa)
T1 / Class1	CG6838-RC	551
T2 / Class2	CG6838-RD	551
T3 / Class2	CG6838-RA	127
T4 / Class1	(Shanks, 2004)	491
T6 / Class1	CG32495-RA	362
T7 / Class1	CG32495-RD	650
T8 / Class2	CG32495-RB	491
T9 / Class3	CG32495-RC	308
T10/ Class2	Isolated in this study	362

The exons have been assigned numbers (Figure 5.1) for easy reference, and the transcripts have been placed in three classes according to their transcriptional start site. These transcript classes differ in the arrangement of their 5′exons (Figure 5.2). Class 1 transcripts contain the

exon designated as exon 1, which is spliced to exon 3. Exon 2 corresponds to the 3′ end of exon 1, intron 1 and exon 3: class 2 transcripts begin with exon 2. All the Class 1 transcripts are predicted to have the start codon in exon 3 Figure 5.2. Class 2 transcripts, containing exon 2, result either from transcriptional initiation within the first intron, or from initiation further to the 5′, without removal of intron 1. Therefore transcripts which start with exon 2 contain an untranslated sequence absent from Class 1 transcripts. However, the translation start codon ATG of the class 2 transcripts is identical to that of the class 1 transcripts (Figure 5.2). All the transcripts except T9, which appears to be a 5′ truncation, share exon 4. All transcripts are identical from exon 9 through to 12, except for T3 which appears to be truncated at the 3′end.

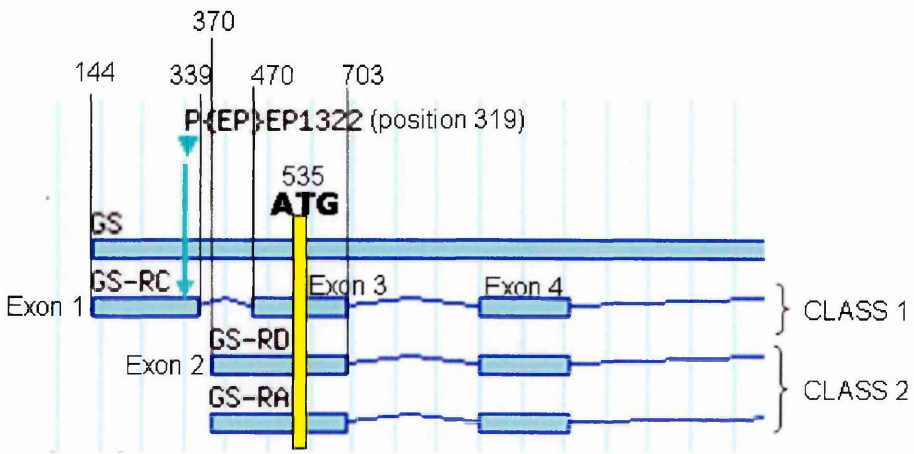


Figure 5.2: Schematic respresentation of the 5′ end of *DmGS*. The positions of the exons, according to the *DmGS* sequence in Appendix B, are shown at the edge of each exon. The ATG start codon is situated at position 535, in the region shared between exons 2 and 3. The P-element EP1322 is situated in the first exon. This schematic diagram is a modification of the FlyBase Genome Browser map (release 4.2.1).

Between exons 4 and 9 are three unique exons named exon 5, exon 7 and exon 8, which can be used to distinguish between the different transcripts. None of these exons correlate to specific transcript classes. The relative positions of each of the exons is listed in Table 5.2. Transcripts T1, T2, T7 and T9 are all predicted to contain exon 5 (Figure 5.1). Because transcript T9 starts in the middle of exon 5, the first exon of T9 was originally named exon 6. It is, however, most likely that T9 is a truncated cDNA clone probably corresponding

to Class 1 or 2, and the exon originally designated exon 6 probably corresponds to a truncated exon5. Exon 7 is thought to exist in T3 and T7, while exon 8 appears in T4, T7 and T8 (Figure 5.1).

Table 5.2: Position of exons, the start codon, 297 element and EP1322 within the *DmGS* genomic region. Positions shown are taken from the transcript alignment against the genomic region in Appendix B.

Genomic element	Position on genomic sequence (Appendix B)
Exon 1	91 – 339
Exon 2	370 – 693
Exon 3	470 – 693
Exon 4	1005 – 1217
Exon 5	2132 – 2478
Exon 7	2996 – 3153
Exon 8	7072 – 7205
Exon 9	7734 – 7776
Exon 10	7922 – 8160
Exon 11	8237 – 8621
Exon 12	8699 – 9001
'297' element	3182 – 6932
ATG start codon	535
EP1322	319

The presence of the nine different predicted transcripts led to the question of which transcripts were present *in vivo*, and whether there was a difference in temporal expression of the transcripts. The first step taken was to test for the existence of each of the unique exons in cDNA isolated from different life stages of *Drosophila*.

5.2.2 Presence of specific exons in Drosophila lifestages

To test whether the exons predicted by FlyBase were present *in vivo*, cDNA was made by RT-PCR from RNA extracted from embryos, first instar larvae, second instar larvae, third instar larvae, pupae, and adult male and female flies. Primers complementary to specific exons were designed in order to test for their presence by PCR amplification. The primers used in this analysis are illustrated in Figure 5.3.

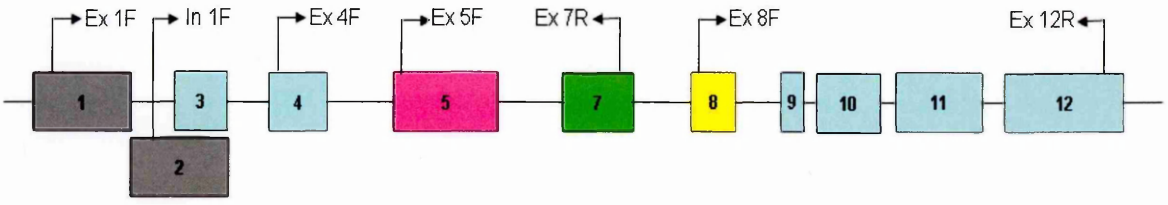


Figure 5.3: Cartoon of the 12 exons present in Drosophila GS, with the positions of primers specific to each of the exons which are labelled with the exon number. The three distinguishing exons are shown in pink, green and yellow boxes. F indicates a forward primer, and R indicates a reverse primer.

The results of this PCR analysis are shown in Figure 5.4 overleaf. Amplification of actin5C was also performed on all the cDNA samples as a positive control. Exon 7 was of particular interest because the only full length transcript containing this exon (T7) was present in Release 3.0 of the *Drosophila* sequence, but was removed from the later Release 4.1 of the sequence.

Primer Pair Ex4F-Ex7R: The first row of Figure 5.4 shows the PCR products using the primer pairs Ex4F and Ex7R, predicted to amplify fragments of 312bp, (exon 4 + exon 7, missing exon 5) or 669bp (exon 4 + exon 5 + exon 7). The size of the smaller band on the gel is consistent with the smaller product, while the larger band is consistent with the larger predicted product. To clarify this issue both the larger and the smaller bands were cloned into TOPO-TA vector and sequenced. The smaller band aligned with the sequence predicted for the exon 4 - exon 7 splice variants. However, the larger band was a non-specific PCR product (corresponding to a region on the 3L chromosome). This result confirms that there are no detectable transcripts in any of the life stages of *Drosophila* which contain exon 5 as well as exon 7, supporting the removal of T7 from the published sequence. It is interesting to note that exon 7 does exist in all life stages, suggesting that T3 may in fact also exist, although whether it is in this truncated form is unclear. The band of intermediate size between the large and small bands observed in the embryo life stage is suspected to be a non-specific PCR artifact since the presence of the band varied with repeated PCR amplification reactions.

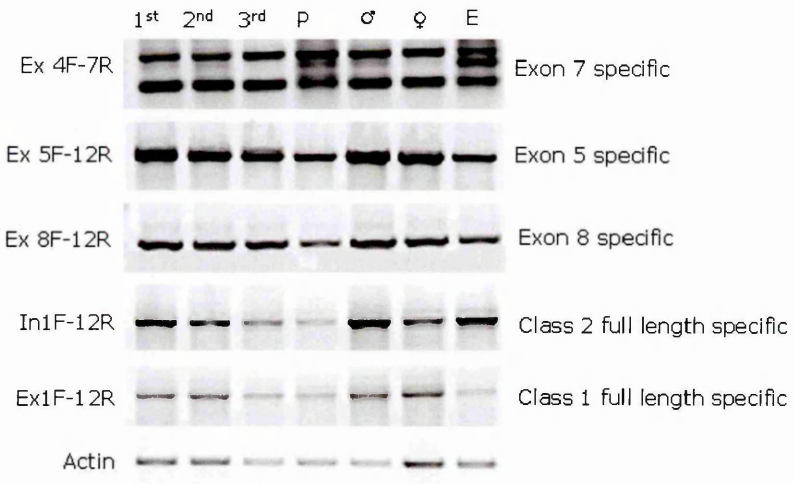


Figure 5.4: PCR fragments from cDNA corresponding to different life stages of *Drosophila*. 1st = 1st instar, 2nd = 2nd instar, 3rd =3rd instar, P = pupa, ♂ = adult male, ♀= adult female, E = embryo. The panels show samples of RT-PCR amplified with primers Ex4F-Ex7R (Exon7 specific), Ex5F-Ex12R (Exon 5 specific), Ex8F-Ex12R (Exon 8 specific), In1F-Ex12R (full-length), Ex1F-Ex12R (full length), and actin5C. The presence of genomic contamination was ruled out because the fragment sizes expected from genomic DNA amplification would have been a larger size than the fragments amplified from cDNA due to the presence of additional intronic sequence in the genomic DNA. There were no additional bands amplified in any of the PCR reactions, indicating the absence of genomic contamination.

Primer pairs Ex5F-12R and Ex8F-12R: The second and the third rows of Figure 5.4 confirm the presence of exon 5 as well as exon 8 in all life stages. The single band when amplifying the cDNA with primer pair ex5F and ex12R confirms the prediction that no transcripts include both exon 5 and exon 8. If there were any further exons between exons 5 and 9 there would have been one or more bands in addition to the band amplified from transcripts lacking any exons between exons 5 and 9 (T1 and T2).

Primer pairs In1F-12R and Ex1F-12R: These primer pairs were designed to amplify the full length transcripts corresponding to Class 1 (Ex1F-12R) and Class 2 (In1F-12R) transcripts. The fourth and fifth rows of Figure 5.4 show amplifications performed with these two primer pairs, with both transcript classes present in all lifestages. Class 1 transcripts, however, appear to be poorly represented in the embryonic stage.

These results show that all the predicted exons are present in all life stages of *Drosophila.* The shortcoming of this approach is that it did not give a complete picture of full length transcripts and the combination of the exons present *in vivo*.

5.2.3 Molecular analysis of DmGS cDNAs

RNA was extracted from male and female flies of the *y sn v* strain shown previously to carry the *DmGS* duplication (Chapter 3), and from *Drosophila* S2 cells which are embryonically derived, and which also contain the *DmGS* duplication. The primer sets used to amplify the whole transcripts (Figure 5.3) were In1F - Ex12R (to amplify transcripts T1, T4, T6, T7), and Ex1F-Ex12R (to amplify transcripts T2, T3, T8, T10). The amplified cDNAs were cloned in the TOPO-TA vector and were then sequenced, as described in Section 2.9 and 2.18.

The results of the sequencing are shown in Figure 5.5, where all the predicted transcripts corresponding to clones sequenced in this study are shown in colour. Transcripts T7 and T9 were not represented in the clones isolated. The absence of T7 was expected, since T7 was removed from the published sequence annotation, and it was verified in the previous section that exon 5 and exon 7could not be detected together in any transcripts. T9 is probably a 5′ truncated cDNA clone, so this experiment would not be expected to identify it as it lacks priming sites for both Ex1F and In1F primers. Similarly transcript T3 would be expected to be absent because it lacks a priming site for primer 12R due to 3′ truncation. Transcript T3 is included in Figure 5.5 despite resting solely on evidence for a transcript containing exon 7 in the previous section - it could not be corroborated in this experiment.

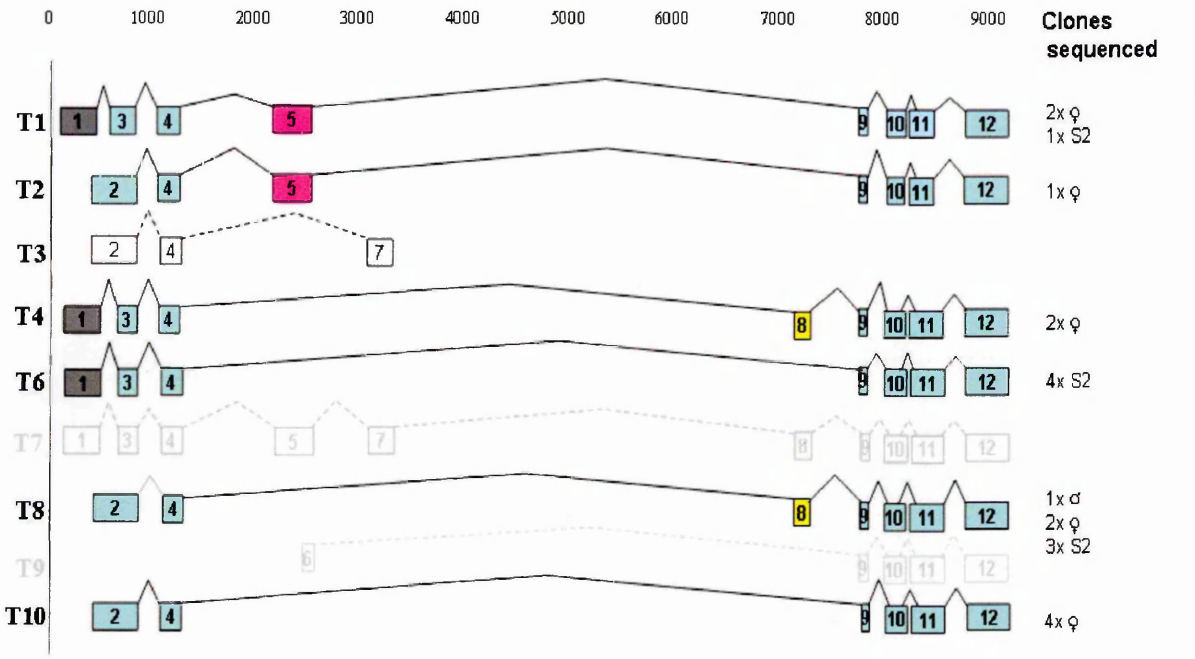


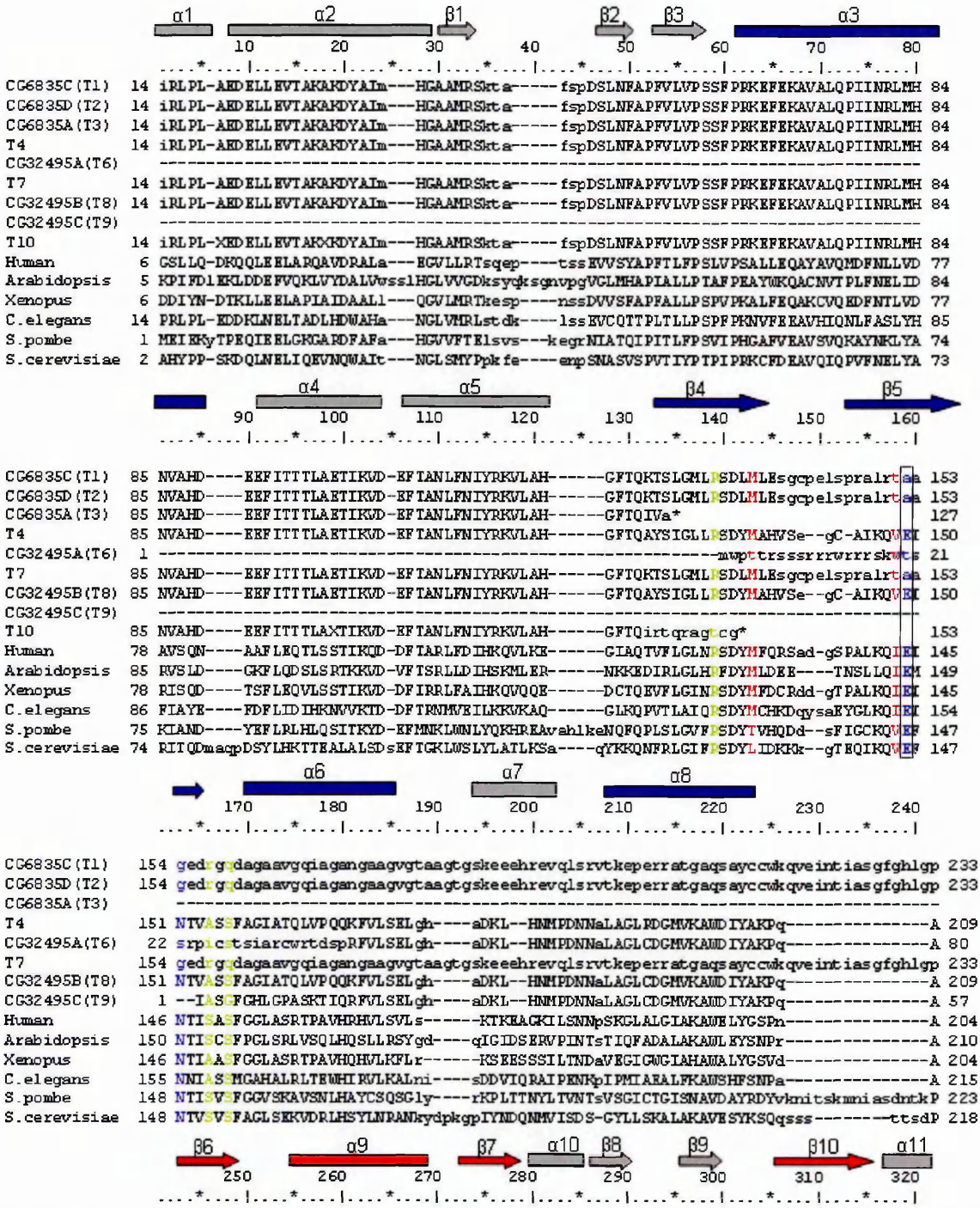
Figure 5.5: Schematic diagram of all the predicted transcripts with information on transcripts present *in vivo*. T3, T7 and T9 are blanked out because there were no clones sequenced which corresponded to these transcripts. The numbers of clones derived from adult fly and S2 cell mRNA that match with each transcript is shown in the column on the right of the diagram.

A new transcript was isolated in this analysis. Transcript T10 has not been previously described, and was found only in female flies. The protein predicted to be encoded by this transcript is identical to that of T6: the only sequence difference between these transcripts being in the 5′ UTR (Figure 5.5). Little can be read into the significance of some transcripts being found only in female flies, or S2 cells because the numbers of clones sequenced is not sufficient for statistical analysis. The high number of transcripts present in female flies is more likely a function of preferential cloning of the cDNAs from that sample.

5.2.4 Bioinformatic analysis of the predicted proteins

Sequencing cDNAs from adult flies and S2 cells confirmed the presence of different *DmGS* transcripts *in vivo*. This result led to further investigation into the function of the different transcripts. A bioinformatic approach of comparing the predicted protein sequences of the *Drosophila GS* transcripts with other eukaryotic GS protein sequences was chosen to gain a

greater understanding of the functions of the *Drosophila* GS proteins. The alignment used the BLAST conserved domain search programme (Marchler-Bauer and Bryant, 2004), and the features of the GS enzyme were annotated using the human protein as the template (Figure 5.6). The protein encoded by each transcript has been named "DrosGS" and the number corresponding to the transcript number, e.g., the protein encoded by T1 is called DrosGS1, and the protein encoded by T2 is called DrosGS2.


α1 α2 β1 β2 β3 α3
α4 α5 β4 β5
α6 α7 α8
β6 α9 β7 α10 β8 β9 β10 α11
CG6835C(T1)
CG6835D(T2)
CG6835A(T3)
T4
CG32495A(T6)
T7
CG32495B(T8)
CG32495C(T9)
T10
Human
Arabidopsis
Xenopus
C.elegans
S.pombe
S.cerevisiae

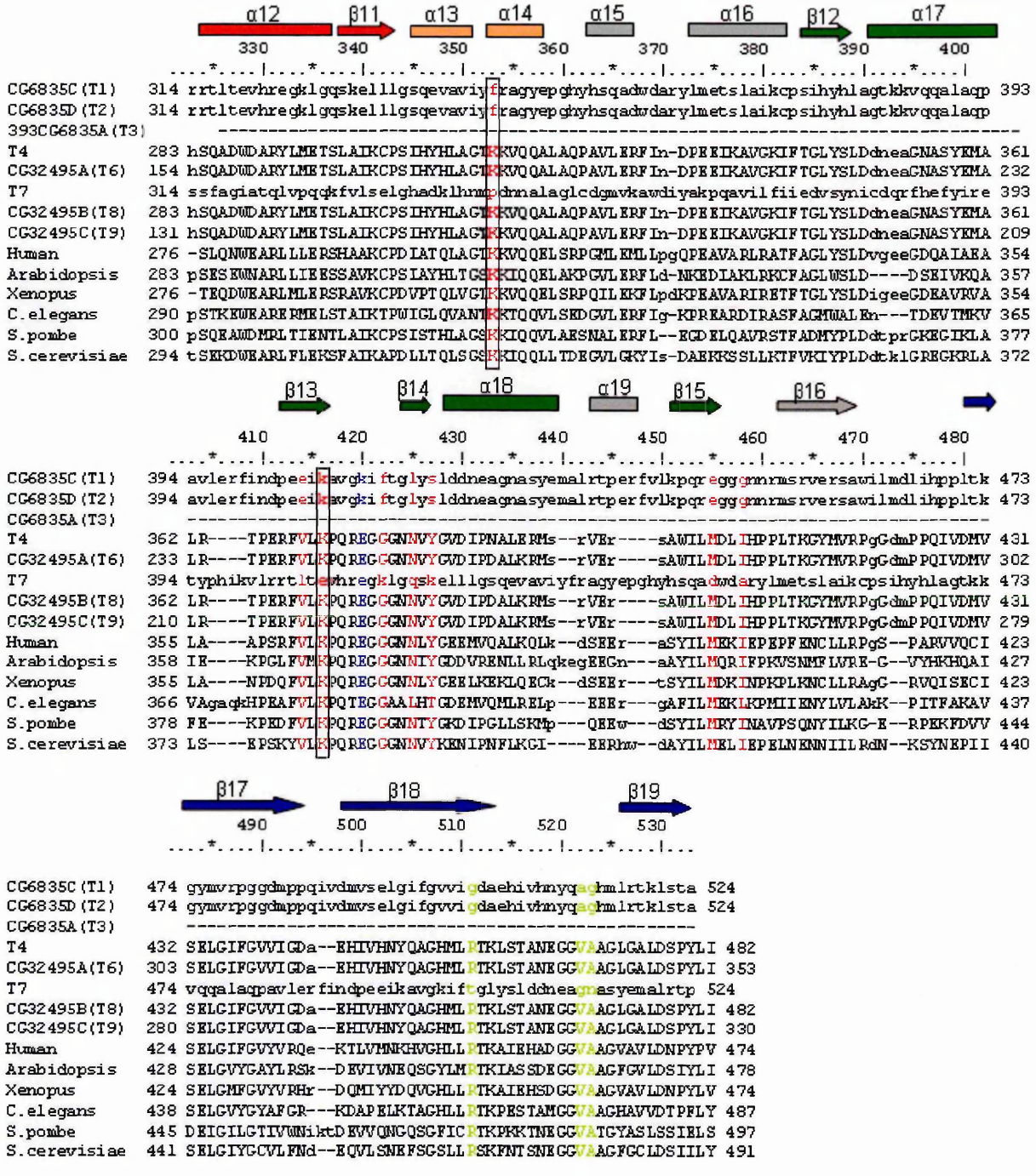


Figure 5.6: ClustalW alignment of the eight *Drosophila* GS protein sequences with representatives of other eukaryotic GS protein sequences created using BioEdit. Active site residues which form specific binding sites in human GS are labelled in green (glutathione binding site), red (ATP binding site) and blue (magnesium binding site). Amino acids that are strictly conserved in virtually all eukaryotic ATP-grasp enzymes are highlighted in boxes (Polekhina *et al.*, 1999). In accordance with Polekhina *et al.* (1999) and Njalsson *et al.* (2004) the domains of the ATP-grasp superfamily are illustrated with coloured bars (α-helixes) and arrows (β-sheets): red, N-terminal domain; orange, linker domain to the lid region; green, lid domain; blue, C-terminal domain. The secondary structures specific to GS are shown with grey coloured bars.

The human GS amino acids conserved in the eukaryotic consensus sequence (Polekhina *et al.*, 1999) are highly conserved in DrosGS4 and DrosGS8. After position Phe-152 (for the sake of consistency, numbering of amino acids in this chapter refers to the human protein), DrosGS6 and DrosGS9 also share a high homology with the consensus eukaryotic sequence for GS (Figure 5.6). Proteins DrosGS1, DrosGS 2, and DrosGS 7 all lose their homology after Met-129, which is caused by a frameshift due to the presence of exon 5 in these proteins. DrosGS3 is terminated prematurely at position 118 on the alignment, followed shortly by the stop codon termination of DrosGS10 at position 126, presumably also caused by a frameshift. Eliminating the proteins that are terminated prematurely, or those that do not share any homology with other eukaryotic GS sequences, DrosGS4, DrosGS6, DrosGS8 and DrosGS9 remain. DrosGS4 and DrosGS8 are the only proteins which share a full homology with eukaryotic GS, while DrosGS6 and DrosGS9 are truncated on the 5′ end, and share homology from residue 152. DrosGS6 is truncated from the 5′end because this protein has a different start codon from the other full length proteins (DrosGS4 and DrosGS8). An alignment of DrosGS6 using the predicted start codon from Flybase, along with the potential DrosGS6 protein initiating from the same ATG as DrosGS8 illustrates how neither start codon produces a full length GS protein from T6 cDNA (Figure 5.7).

Figure 5.7: Protein alignment of DrosGS6 with two alternative start codons and DrosGS8. DrosGS6-FB is the predicted protein sequence on Flybase which corresponds to T6. DrosGS6-st is the predicted protein from the T6 cDNA sequence using the same start codon as DrosGS8. The residues shared between DrosGS6-FB and DrosGS8 are shown in blue, while the residues shared between DrosGS6-st and DrosGS8 are shown in pink.

The critical amino acids in the human GS catalytic site are highly conserved in these four remaining GS proteins (note that DrosGS6 and DrosGS9 only start at residue 152). The residues that bind glutathione are conserved (Arg-124, Ser-151, Asn-216, Gln-220, Arg-267, Tyr-270, Arg-450, Val-461, Ala-462), except for residues **Ser-149, Glu-214** (Table 5.3). Human Ser-149 is an alanine residue in DrosGS4 and DrosGS8. Although Ser-149 is the consensus sequence at this position, *Xenopus* and *C. elegans* proteins share an alanine at that position with *Drosophila*. Glu-214 is a serine residue in the *Drosophila* sequence, and it appears that this is a variable residue in two of the other eukaryotic sequences as well.

The residues that form the magnesium ion binding sites (Glu-144, Asn-146, Glu-368) are conserved in DrosGS4 and DrosGS8, however in DrosGS6 and DrosGS9, the first two residues listed are not present due to the truncated proteins. The residues contributing to the

ATP binding site are also conserved (Met-129 Lys-305, Val-362, Lys-364, Gly-370, Asn-373, Tyr-375, Met-398, Ile-401) except for **Ile-143**, which is a valine residue in the *Drosophila* proteins (Table 5.3). The Ile-143 residue is however variable amongst eukaryotes, and the *Drosophila* valine is shared by *S. cerevisiae* as well as *S. pombe* at this position. The three residues that differ from the human sequence (Ser-149, Glu-214 and Ile-143) are all conservative substitutions, and therefore should not alter the protein function dramatically (Betts and Russell, 2003).

The three residues that are strictly conserved (Polekhina et al., 1999; Dinescu et al., 2004) throughout all ATP grasp family members (Lys-305, Lys-364, and Glu-144) are also conserved in DrosGS4 and DrosGS8, as well as in DrosGS6 and DrosGS9, although in the latter two the Glu-144 residue is in the truncated portion of the protein.

These results strongly suggest that the only active GS protein is likely to be translated from transcripts T4 and T8. The function of the other transcripts which, on the basis of this sequence comparison, do not appear likely to encode an active GS protein but which are present *in vivo*, is unknown at this time. In order to confirm the biological activity of DrosGS4/DrosGS8, the T8 cDNA was used in a yeast complementation experiment.

Table 5.3: Position of conserved GS residues involved in critical GS functions. The residue positions are relative to the human GS protein sequence. The four *Drosophila* GS proteins listed are the full length DrosGS4 and DrosGS8, as well as the truncated proteins DrosGS6 and DrosGS9. Conserved residues shared with the human protein are shaded.

	Residue position	Human	DrosGS4	DrosGS6	DrosGS8	DrosGS9
GS catalytic site	124	Arg	Arg	Arg	Arg	Arg
	149	Ser	Ala	---	Ala	---
	151	Ser	Ser	Ser	Ser	Ser
	214	Glu	Ser	Ser	Ser	Ser
	216	Asn	Asn	Asn	Asn	Asn
	220	Gln	Gln	Gln	Gln	Gln
	267	Arg	Arg	Arg	Arg	Arg
	270	Tyr	Tyr	Tyr	Tyr	Tyr
	450	Arg	Arg	Arg	Arg	Arg
	461	Val	Val	Val	Val	Val
	462	Ala	Ala	Ala	Ala	Ala
ATP binding	129	Met	Met	Met	Met	Met
	143	Ile	Val	Val	Val	Val
	305	Lys	Lys	Lys	Lys	Lys
	362	Val	Val	Val	Val	Val
	364	Lys	Lys	Lys	Lys	Lys
	370	Gly	Gly	Gly	Gly	Gly
	373	Asn	Asn	Asn	Asn	Asn
	375	Tyr	Tyr	Tyr	Tyr	Tyr
	398	Met	Met	Met	Met	Met
	401	Ile	Ile	Ile	Ile	Ile
Mg binding	144	Glu	---	Glu	---	Glu
	146	Asn	---	Asn	---	Asn
	368	Glu	Glu	Glu	Glu	Glu

5.2.5 *Complementation of S. pombe deficient in GS by a Drosophila cDNA*

5.2.5.1 GS deficient *S. pombe* strain MN101

Schizosaccharomyces pombe was chosen for functional complementation experiments because it is a eukaryotic system which has proved useful for expression of heterologous proteins.

MN101 is a GS null *S. pombe* strain derived from the parental strain HM123 (h^-, leu^{-1}), (Mutoh and Hayashi, 1988). Loss of GS in MN101 results in a cadmium hypersensitivity because glutathione synthesising enzymes are involved in the formation of phytochelatin, a molecule that forms complexes with cadmium (Mutoh and Hayashi, 1988). Cadmium is highly toxic to cells, and is detoxified in plants and yeast by phytochelatin. Biosynthesis of phytochelatin [$(\gamma$-Glu-Cys$]_n$-Gly] is controlled by GCL and GS (Al-Lahham et al., 1999), the same enzymes involved in glutathione production, as described in Chapter 1. Cadmium is also known to cause glutathione depletion, thereby resulting in increased oxidative stress (Radyuk *et al.*, 2003). Wild type *S. pombe* strain HM123 has slightly inhibited growth in the presence of 0.5mM $CdCl_2$, while the growth of MN101 mutants is severely inhibited at this concentration, as shown in Figure 5.8. Strain MN101 has no GS activity (Mutoh et al., 1991), and the minimal amounts of glutathione measured in the cells (less than 17% of wild type) is most likely derived from trace GSH in the media (Mutoh and Hayashi, 1988).

The cadmium hypersensitivity of strain MN101 can be used in an assay for testing the activity of different transgenic GS proteins. Functional complementation of MN101 GS has been achieved with GS cDNAs from the mouse (Shi *et al.*, 1996) as well as from *Arabidopsis thaliana* (Wang and Oliver, 1996).

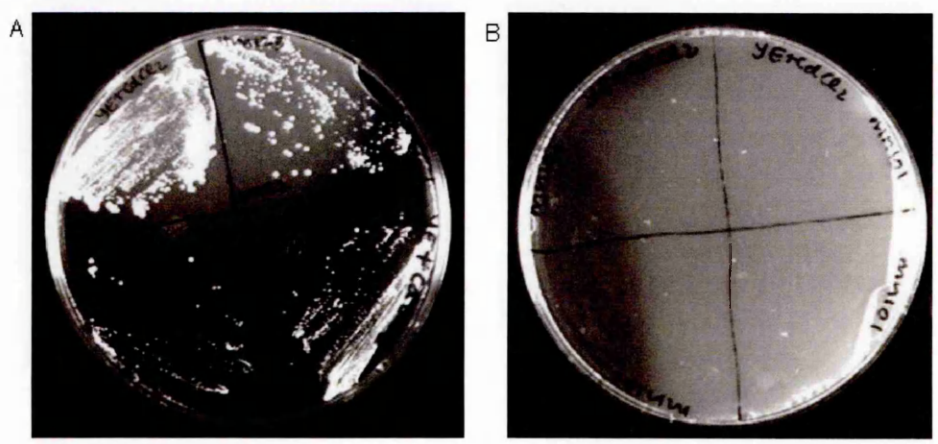


Figure 5.8: Wild type parent strain HM123 (Panel A) and the GS deficient MN101 strain which is sensitive to $CdCl_2$ (Panel B). Each of the strains was grown on YE plates containing 0.5mM $CdCl_2$.

5.2.5.2 Complementation of *S. pombe* deficient in GS

The bioinformatic analysis performed on the *DmGS* transcripts in the previous section suggested that the GS protein predicted to be encoded by transcripts T4 and T8 would be the active protein. To test the biological function of the protein encoded by these two cDNAs, a complementation experiment was performed using a *S. pombe* strain lacking GS. The protein encoded by T4 and T8 is predicted to be identical, therefore a cDNA corresponding to T8 was cloned in the plasmid pNMT1-TOPO for further expression studies. The resulting clone was named MN101_10. Because T6 had strong homology with the eukaryotic GS consensus (apart from the 5′ truncation), a similar construct, MN101_12, containing this cDNA was also used in this experiment. T8 and T6 correspond to the sequences listed on FlyBase as CG32495-B and CG32495-A respectively. *S. pombe* mutant strain MN101 was transformed with each of these constructs (Figure 5.9).

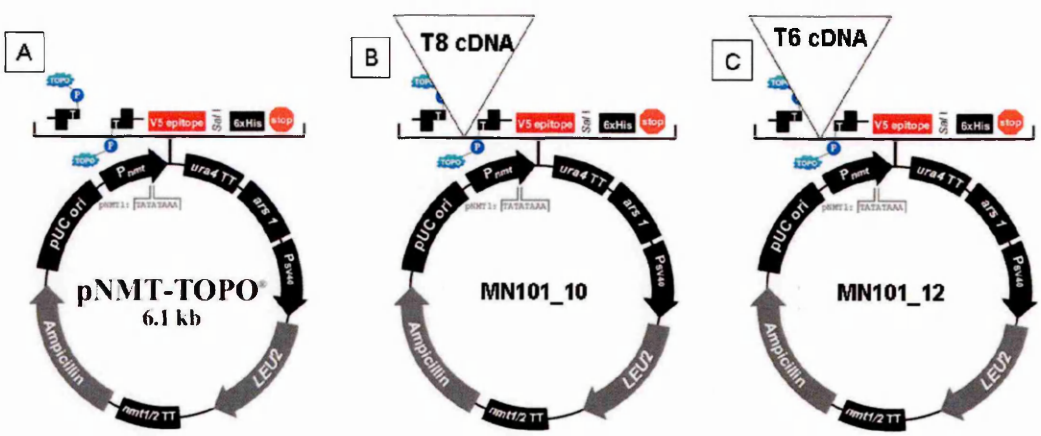


Figure 5.9: Constructs using the pNMT1-TOPO vector. Panel A: Schematic diagram of the features of the pNMT1-TOPO vector. Figure was obtained from www.invotrogen.com. Panel B: Construct MN101_10 containing cDNA from transcript T8 (CG32495-B). Panel C: Contruct MN101_12 containing cDNA from transcript T6 (CG32495-A).

The cDNAs were amplified using PCR, and cloned in the TOPO vector pNMT1-TOPO. The pNMT1-TOPO vector illustrated in Figure 5.9 contains the *S. pombe ars1* origin of replication which maintains a high copy number of the plasmid in *S. pombe* cells, and does not require chromosomal integration. The selection marker for *E. coli* cells is ampicillin resistance, and in

yeast cells it is a *S. cerevisiae LEU2* auxotrophic marker that allows growth of transformed cells on medium that does not contain leucine. The transgene expression is driven by an *nmt1* promoter which allows for thiamine-regulated expression in *S. pombe* cells by switching off gene expression in the presence of 15μM thiamine. This promoter has, however, been found to still allow gene expression in the presence of thiamine, and the ability to 'shut off' completely seems to be gene-dependent (Forsburg, 1993). For this reason the thiamine control was not used in this experiment. The NMT1 version of the promoter used in these complementation experiments offers high expression levels of protein expression (Basi *et al.*, 1993). The plasmid also has a C-terminal V5 epitope and a polyhistidine tag, however neither was used in this experiment as the cDNA sequences cloned into the plasmid contained the native stop codon at the 3′ end of exon 12.

T6 and T8 clones were confirmed, and the orientation of the insert verified, by sequencing using the primers NMT (forward) and URA4 (reverse). MN101 cells were transformed with verified clones as described in Section 2.11. As shown in Figure 5.10, the GS protein encoded by T8 (clone MN101_10) was able to complement the loss of GS function in MN101. The heterologous GS allowed the yeast to produce phytochelatins, thereby restoring the ability of the mutant MN101 to grow in the presence of 0.5mM $CdCl_2$, although colony growth was slower than the wild type strain HM123. The product of transcript T6 (clone MN101_12) did not, however, rescue the cadmium hypersensitivity of MN101.

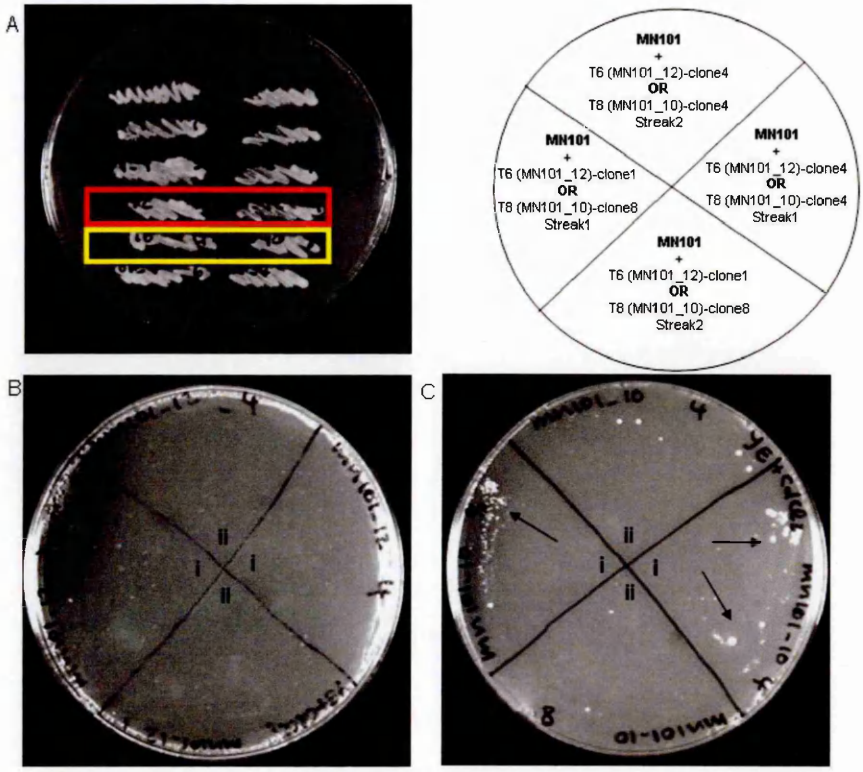


Figure 5.10: *S. pombe* GS deficient MN101 strain transformed with two different *GS* cDNAs. Panel A shows all the *S. pombe* clones transformed with different *Drosophila GS* cDNAs. The two MN101 clones carrying pNMT-TOPO_10 (transcript T8) are highlighted in red, and two clones carrying pNMT-TOPO_12 (transcript T6) are highlighted in yellow. The clones were grown on the plasmid selection medium (EMM without Leu). In Panels B and C, MN101 cells containing a pNMT-TOPO plasmid with a *GS* transcript were grown on YE plates containing 0.5mM $CdCl_2$. Two different clones containing the same transcript were tested (Colony 4 and 8 for MN101_10, and Colony 1 and 4 for MN101_12), and each different clone was streaked out twice (quarters of the plates labelled i and ii), to allow for even colony spread. Panel B shows MN101 cells transformed with a pNMT-TOPO containing transcript T6. This cDNA did not result in expression of a protein which could rescue the GS deficiency. Panel C shows MN101 cells transformed with a pNMT-TOPO containing transcript T8. The protein expressed in these cells did rescue the GS deficiency in the MN101 *S. pombe* strain, although the colonies were slow growing (colonies are indicated with arrows).

To test whether the *S. pombe* cells containing the MN101_10 or MN101_12 constructs were expressing a transgenic protein, SDS-PAGE was carried out on soluble protein extracted from the yeast cells. Coomassie staining of the separated proteins (Figure 5.11A) revealed no distinct band in the yeast cells transformed with MN101_10 or MN101_12 compared to MN101 with no plasmid, and the parental strain HM123. It was expected that there would be a unique 62 kDa band in MN101_10 cells which carried T8 cDNA, and another distinct band of approximately 52 kDa in MN101_12 cells carrying T6 cDNA, as this is the predicted size of the truncated protein coded for by T6. It was not possible to distinguish any unique bands,

which could be because this method of staining is not sensitive enough to detect one additional band. A Western blot of the separated yeast proteins probed with an antibody to *Drosophila* GS (a gift from Lesley McLellan) is illustrated in Figure 5.11B. The anti-DmGS antiserum was raised against the T2 protein (Shanks, 2004).

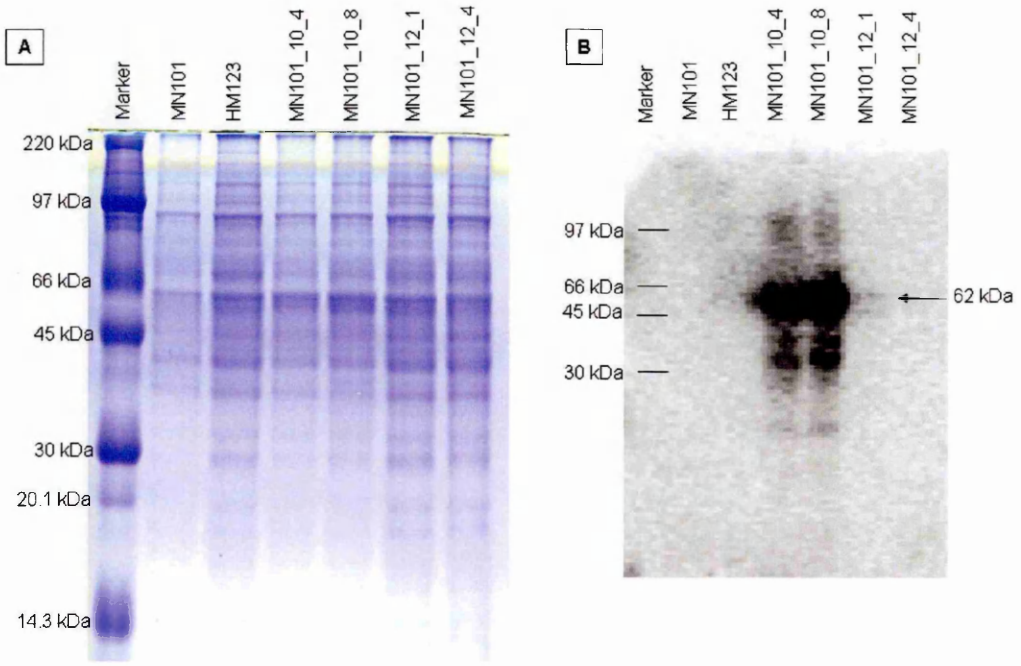


Figure 5.11: SDS-PAGE separation of proteins extracted from *S. pombe* GS mutant strain MN101 and HM123 parental strain, as well as MN101 carrying pNMT-TOPO_10 (transcript T8) and pNMT-TOPO_12 (transcript T6). MN101_10_4 and MN101_10_8 refer to two clones carrying transcript T8, and MN101_12_1 and MN101_12_4 refer to two clones carrying transcript T6. Panel A: A Coomassie Blue stained SDS-PAGE gel. A total of 20µg protein was loaded in each lane. Coomassie staining detects bands as low as 50-100ng. Panel B: Western blot of the same gel as in Panel A. Only the *S. pombe* cells carrying pNMT-TOPO_10 expressed a 62kDa protein recognised by a *Drosophila* GS antibody.

Drosophila GS was only detected in yeast cells carrying the MN101_10 construct. The parental HM123 and MN101 strains were not expected to have a positive signal for *Drosophila* GS, however it is surprising that MN101_12 containing cells did not have any reaction with the *Drosophila* antiserum. MN101 cells transformed with the construct containing T6 cDNA either do not synthesise the predicted DrosGS protein, or the truncated protein they create does not bear the epitopes detected by the *Drosophila* GS antisera, which was raised to transcript T2 protein (Shanks, 2004). It is unknown at this stage what epitopes

are recognised by this *Drosophila* GS antibody, but it is possible that they lie in the N terminal region of the protein truncated in the protein encoded by T6. It is unfortunate that the predicted truncated protein could not be distinguished by the Coomassie stain.

5.2.6 DmGS expression patterns in Drosophila

To determine the temporal and spatial expression patterns of *DmGS* in *Drosophila*, rtPCR was used to compare the level of total *DmGS* transcripts in different life stages of *Drosophila*, as well as in different tissues of the adult fly. The primer pair used to amplify total *DmGS* transcripts was 'first exon_F' and 'Ex12_R'. This primer pair spans all *DmGS* transcripts, with the forward primer situated in the region shared by exon 1 and intron 1, and the reverse primer situated on the shared terminal exon (Figure 5.3). To assess the relative levels of *DmGS* transcripts, a semi-quantitative PCR approach was taken. Amplification of actin5C was used as an internal reference against which levels of *DmGS* were estimated, using the Typhoon 9410 variable mode imager and Imagequant software, as described in Section 2.17. Each experiment was performed at least three times to allow for statistical analysis of the results.

5.2.6.1 Relative expression of *DmGS* in different life stages

Larval full length *DmGS* transcript levels were highest in the first instar, and reduced steadily as the larvae grew, reaching the lowest levels in the pupal phase (Figure 5.12). *DmGS* expression levels in adult male and female flies was comparable to first and second instar larvae. *DmGS* transcript levels are minimal during embryonic development. During first larval instar, the feeding rate is known to increase, reaching a peak at the transition between the first and second instar, after which the rate declines to the lowest point in the third instar (Ruiz-Dubreuil et al., 1996; Godoy-Herrera et al., 2004). It is possible that this reduction in feeding rate in second, and especially third, instar larvae creates less of an antioxidant requirement compared to the first instar stage when the larvae are exposed to a wide range of

environmental oxidants through the diet. It is unclear why male flies have a higher *DmGS* expression level than the female flies, although this sex specific difference will be analyzed in detail in the following chapter.

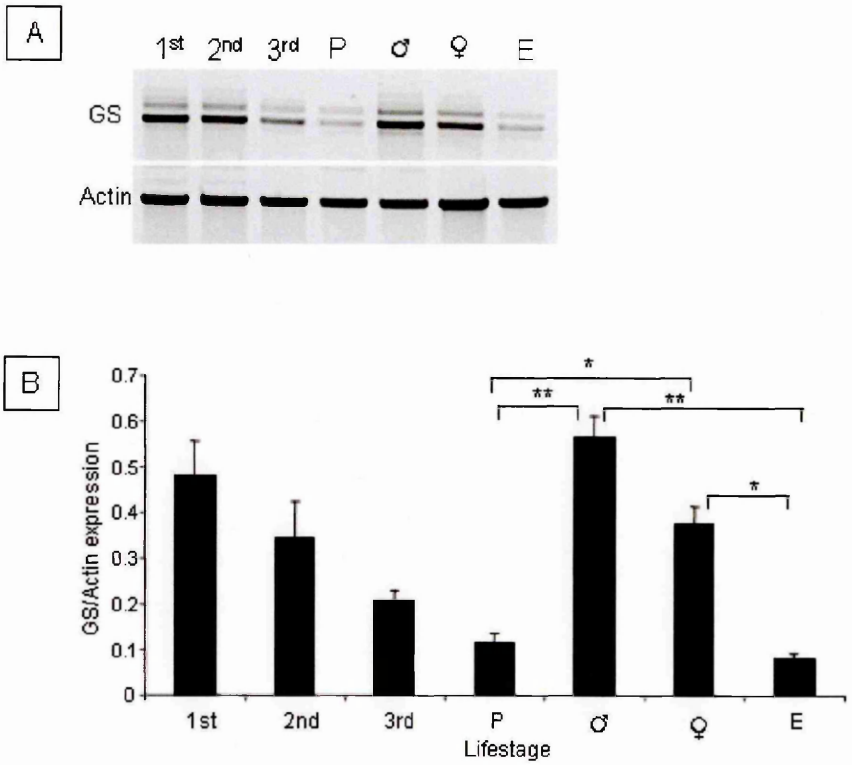


Figure 5.12: *GS* expression in different life stages of the single copy *GS* duplication reference strain (*w sn*^{3}). The lifestages shown are 1st instar (1st), 2nd instar (2nd), 3rd instar (3rd), pupa (P), male adult (♂), female adult (♀) and embryo (E). Panel A shows samples of GS and actin RT-PCR separated on an agarose gel. Panel B is a graphical representation of the relative intensities of the GS:Actin RT-PCR product from the gel in Panel A. Means±SEM are shown. Statistical differences between GS transcript levels in each of the consecutive lifestages was determined by Student's T-test, where *=p<0.05 and **=p<0.01.

In all life stages, the smaller *DmGS* band, corresponding to transcript T4/T8 based on size, is denser than the larger band, which corresponds to T1/T2 (Figure 5.12). This suggests that the T4/T8 transcripts, which are highly likely to encode the active GS protein, are significantly more abundant than the T1/T2 transcripts in all *Drosophila* developmental stages.

5.2.6.2 Relative expression of *DmGS* in different adult fly body parts

To determine whether *DmGS* was more highly expressed in one particular tissue type in *Drosophila,* RNA was extracted from different regions of the adult fly. Because of the small

size of *Drosophila*, preparation of purified tissue would be very demanding, therefore body parts enriched for specialised tissues were used (Zheng et al., 2005; Girardot et al., 2006). The head was used to represent neural tissue, the thorax represented muscle tissue, and the abdomen represented reproductive tissue. The level of full length *DmGS* transcripts in different tissues was determined in male as well as female flies.

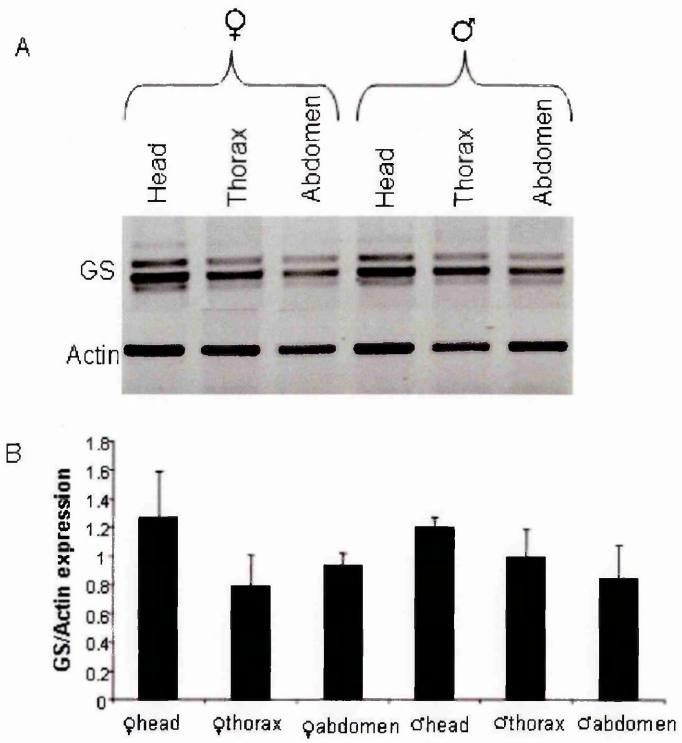


Figure 5.13: *DmGS* expression in different body parts of adult males and females. The analysis was performed on the *DmGS* duplication reference strain (*y sn v*). Panel A shows samples of *DmGS* and actin RT-PCR PCR separated on an agarose gel. Panel B is a graphical representation of the GS:Actin relative intensities from the gel in Panel A. There was no significant difference in GS expression between different body parts in males or females. Statistical significance was calculated using Students 't'test. Means±SEM are shown.

5.3 Discussion

The results presented in this chapter show that the numerous transcripts predicted for *DmGS* do in fact exist. The only other instance of multiple *GS* transcripts in the literature is in the mouse, where six transcripts encode three different proteins (Shi *et al.*, 1996). The *Drosophila GS* gene is predicted to produce nine transcripts, six of which have been found *in vivo* in this study. It is possible that the remaining three transcripts also exist *in vivo*, however in the case

of transcript T7, no evidence was found for exons five and seven existing together in the same transcript. The remaining two transcripts, T3 and T9, are both truncated transcripts, and although they may have a biological function, it is more likely that they exist merely as cloning or sequencing artefacts of incomplete EST clones. It is unclear as to how the different transcripts were assigned to each of the two copies of *DmGS* because the sequence of the two copies is identical except for the *297* element present in the distal copy.

The predicted protein sizes from each of the *Drosophila* transcripts are more varied than any of the eukaryotic GS sequences published to date, which are in the size range of 474 to 492 amino acids, as listed in Table 5.4. The two *Drosophila* transcripts which encode a protein in this size range (491aa) are T4 and T8, and T8 cDNA was able to rescue the GS deficient *S. pombe* strain MN101.

Table 5.4: Comparison of the sizes of eukaryotic GS proteins

Organism	GS protein size (amino acids)	Reference
Homo sapiens	474	(Njalsson *et al.*, 2004)
Arabidopsis thaliana	478	(Wang and Oliver, 1997)
Schizosaccharomyces pombe	498	(Phlippen *et al.*, 2003)
Saccharomyces cerevisiae	492	(Grant *et al.*, 1997)
Mus musculus	474 (active protein) & 432	(Shi *et al.*, 1996)
Rattus norvegicus	474	(Huang *et al.*, 1995)
Drosophila melanogaster	491 (predicted active)	Predicted protein size

A comparison of the *Drosophila* amino acid sequence with other eukaryotic GS sequences revealed a high homology of *Drosophila* GS (DrosGS4 and DrosGS8) to the other eukaryotic sequences. Only three (Ser-149, Glu-214, Ile-143) of the 24 active site residues identified in human GS were different in *Drosophila* GS, and none of these three residues are one of the 16 mutations known to alter GS activity (Polekhina et al., 1999; Njalsson et al., 2004). Neither Ser-149 nor Glu-214 appear to play a role in the secondary structure of GS, while Ile-143 is situated in a N-terminal β-sheet named β5 in Figure 5.6. Because the only residues that vary

between DrosGS4/DrosGS8, and the human protein at the active site are three residues which are not strictly conserved across eukaryotic GS sequences, it is likely that the protein from these two transcripts is the active GS protein in *Drosophila.*

DrosGS6 and DrosGS9 share a close homology with other eukaryotic GS protein sequences, however they have a large portion of the N-terminal end of the sequence missing, thereby losing the first six of the twenty four active site residues. Included in those six residues are three of the four residues that create the magnesium binding site. It is therefore unlikely that the proteins encoded by transcripts T6 and T9 are active GS enzymes. Deletions of the 5′ region of the *Arabidopsis GS* gene resulted in complete loss of enzyme activity, even when only 5 amino acids were deleted (Wang and Oliver, 1997). This suggests an essential function for the N-terminal end of *GS* in determining enzyme activity, and casts more doubt on the likelihood of transcripts T6 and T9 encoding active proteins.

The structure of the transcript which encodes the active form of GS in *Drosophila* appears to be that of transcripts T4 and T8: after the start codon, the presence of exons 4, 8, 9, 10, 11, and 12 result in a RNA transcript which codes for a GS protein with a high sequence homology to other eukaryotic GS sequences, and shares 21 of the 24 active site residues with human GS. Transcripts T4 and T8 are predicted to encode identical proteins, but have different transcription start sites. Transcripts T1, T2 and T7 lose homology with the eukaryotic GS sequence at the start of exon 5, while T3 and T10 are truncated by a stop codon in exon 7.

The presence of transcripts which do not produce an active protein has also been noted in mice, and the distribution of the different murine transcripts was tissue specific (Shi *et al.*, 1996). In both *Drosophila* and mice it appears that only one active protein is produced from the multiple *DmGS* transcripts present in the cell.

The biological activity of transcripts T4 and T8 was confirmed by complementation of the GS deficient *S. pombe* strain MN101. Gene complementation in *S. pombe* was chosen because it is a eukaryotic system which has proved useful for expression of heterologous proteins (Forsburg, 1993). The MN101 GS deficient strain has been used for functional complementation of GS in the mouse (Shi *et al.*, 1996) as well as in *Arabidopsis* (Wang and Oliver, 1996).

The *S. pombe* GS protein contains 498 amino acids and has an estimated molecular mass of 56.2 kDa (Wang and Oliver, 1997), while the *Drosophila* protein is 491aa long, and has a MW of 62 kDa. The GS proteins from each species share only a 36% homology, as illustrated in Figure 5.14. Despite the differences between the *S. pombe* and *Drosophila* protein sequences, the T8 *Drosophila* cDNA complements MN101 GS deficiency by encoding a GS protein that restored glutathione synthesis in the mutant MN101. This in turn is assumed to have increased the phytochelatin levels in MN101, and restored cadmium resistance. Glutathione was probably not restored to wild type levels as growth on cadmium was slower than the wild type HM123 strain. This lack of complete activity has been noted previously with heterologous expression of GS synthesis proteins (Ullmann et al., 1996; Saunders and McLellan, 2000), and may be because the *Drosophila* protein requires factors not present in *S. pombe* for full activity. *S. pombe* contains a 15aa sequence between helix α8 and beta sheet β6 which does not occur in any other eukaryotic sequences, as shown in Figure 5.14 (Phlippen *et al.*, 2003) . It is feasible that the presence of that unique sequence is necessary for optimal functioning within *S. pombe* cells.

```
S. pombe   9    EQIEELGKGARDFAFAHGVVF-TELSVSKEGRNIATQIPITLFPSVIPHGAFVEAVSVQK  67
                +++ E+   A+D+A  HG    ++ + S +  N A   P  L PS  P   F +AV++Q
D. melan  21    DELLEVTAKAKDYAIMHGAAMRSKTAFSPDSLNFA---PFVLVPSSFPRKEFEKAVALQP  77

S. pombe  68    AYNKLYAKIANDYEFLRLHLQSITKYDEFMNKLWNLYQKHREAVAHLKENQFQPLSLGVF  127
                  N+L   +A+D EF+   L    K DEF   L+N+Y+K    +AH      Q  S+G+
D. melan  78    IINRLMHNVAHDEEFITTTLAETIKVDEFTANLFNIYRK---VLAH---GFTQAYSIGLL  131

S. pombe  128   RSDYMVHQDDSFIGC--KQVEFNTISVSFGGVS-KAVSNLHAYCSQSGLYRKPLTTNYLT  184
                RSDYM H  +   GC  KQVE NT++ SF G++ + V       S+ G   K    + +
D. melan  132   RSDYMAHVSE---GCAIKQVEINTVASSFAGIATQLVPQQKFVLSELGHADK---LHNMP  185

S. pombe  185   VNTSVSGICTGISNAVDAYRDYVKNITSKMNIASDNTKPIVLFVVKGGERNITDQRTLEY  244
                 N +++G+C G+  A D Y                  + ++LF+++    NI DQR  E+
D. melan  186   DNNALAGLCDGMVKAWDIYA---------------KPQAVILFIIEDVSYNICDQRFHEF  230

S. pombe  245   ELLNRF-HVISKRIDIAELNSLIHDKSSNKLYMKTSFTTYEVAVVYYRVGYALDDYPSQE  303
                +   + H+   R  + E++       S +L + +     EVAV+Y+R GY    Y SQ
D. melan  231   YIRETYPHIKVLRRTLTEVHREGKLGQSKELLLGSQ----EVAVIYFRAGYEPGHYHSQA  286

S. pombe  304   AWDMRLTIENTLAIKCPSISTHLAGSKKIQQVLAESNALERFL-EGDELQAVRSTFADMY  362
                 WD R  +E +LAIKCPSI  HLAG+KK+QQ LA+   LERF+ + +E++AV   F  +Y
D. melan  287   DWDARYLMETSLAIKCPSIHYHLAGTKKVQQALAQPAVLERFINDPEEIKAVGKIFTGLY  346

S. pombe  363   PLDDTPRGKEGIKLAFEKPEDFVLKPQREGGGNNTYGKDIPGLLSKMPQEEWDSYILMRY  422
                 LDD   G    ++A   PE FVLKPQREGGGNN YG DIP  L +M + E  ++ILM
D. melan  347   SLDDNEAGNASYEMALRTPERFVLKPQREGGGNNVYGVDIPDALKRMSRVERSAWILMDL  406

S. pombe  423   INAVPSQNYILK--GERPEKF-DVVDEIGILGTIVWNIKTDEVVQNGQSGFICRTKPKKT  479
                I+   ++ Y+++  G+ P +  D+V E+GI G ++ +   + +V N Q+G + RTK
D. melan  407   IHPPLTKGYMVRPGGDMPPQIVDMVSELGIFGVVIGD--AEHIVHNYQAGHMLRTKLSTA  464

S. pombe  480   NEGGVATGYASLSS  493
                NEGGVA G  +L S
D. melan  465   NEGGVAAGLGALDS  478
```

Figure 5.14: Alignment of *S. pombe* and *D. melanogaster* GS protein sequences using BLAST 2 (Tatusova and Madden, 1999).

T4/T8 transcripts are significantly more abundant than the T1/T2 transcripts in all *Drosophila* developmental stages. This observation fits the hypothesis that the protein encoded by transcripts T4/T8 is the active GS protein in *Drosophila*. Different levels of *GS* expression occur in mice and rats, with the majority of enzyme activity being located in the kidneys, followed by the liver (Meister, 1985; Shi *et al.*, 1996). In *Drosophila* neural tissue specific overexpression of antioxidant genes has extended lifespan (Parkes et al., 1998; Ruan et al., 2002; Morrow et al., 2004; Bauer et al., 2005), raising the question as to whether there would be a particular spatial expression pattern of *GS*. The lack of tissue-specific *DmGS* expression observed may suggest that GS is an essential enzyme required throughout the organism by all tissue types. In this study, the temporal expression pattern of *DmGS* transcripts showed that

GS is ubiquitous throughout all developmental stages, but with stronger expression in early larval instars and adults.

The question arising from these data is: What is the function of these multiple transcripts? It can be speculated that the RNA transcripts or even the proteins produced by the different transcripts are involved in GS regulation. The different RNA transcripts may also be regulated in different ways, thereby contributing to the regulation of GSH in different organs. It remains to be investigated whether the transcripts besides T4/T8, or their putative proteins, are involved in GS regulation.

6 Chapter 6: Role of *GS* in oxidative stress resistance of adult *Drosophila*

6.1 Introduction

Glutathione (GSH) has many cellular functions, including protecting cells against oxidative damage, amino acid transport, detoxification of xenobiotic compounds, provision of reducing equivalents, and being a co-factor for many enzymes (Kehrer and Lund, 1994; Meister, 1994; Polekhina et al., 1999). In its role as antioxidant, GSH can react directly with hydroxyl radicals ($OH^{\bullet}$) reducing them to H_2O, and it also serves as an electron donor to glutathione peroxidase which detoxifies peroxides (H_2O_2) (Inoue et al., 1998; Parkinson, 2001). The antioxidant activity of GSH is of interest in ageing studies as there is evidence that raising the antioxidant defences of an organism can result in increased oxidative stress resistance, and an increase in longevity (Finkel and Holbrook, 2000; Sohal et al., 2002).

The correlation between oxidative stress resistance and longevity in many organisms including worms, flies and mice has been well documented (Wallace, 1999; Finkel and Holbrook, 2000; Hekimi and Guarente, 2003). Because oxidative stress resistance and longevity are closely linked, screening for factors that influence oxidative stress resistance is an effective alternative to the more time consuming direct screening for lifespan in studies (Lin et al., 1998).

The capacity of cells to regenerate glutathione during an oxidative challenge is important in preventing loss of cellular function due to oxidative damage to cellular components, and *de novo* GSH synthesis plays an important role in maintaining cellular GSH levels (Shi et al., 1994). GSH is synthesised in a two-step reaction with the first rate limiting step catalysed by glutamate-cysteine ligase (GCL) and the second step is catalysed by glutathione synthetase (GS). GCL is a heterodimer composed of a catalytic subunit (GCLC) and a regulatory subunit called GCLM. GCLC has catalytic function independent

of GCLM, however the binding of GCLC to GCLM greatly increases the activity of GCL by increasing the affinity of the enzyme for the substrates glutamate and cysteine (Huang et al., 1993a; Huang et al., 1993b). *Drosophila* GCLC, which shares the same catalytic properties as human and rat GCLC (Fraser et al., 2002), has been overexpressed in the motorneurons of flies, with the result of increasing the glutathione content, longevity and resistance to oxidative stress in transgenic flies (Orr et al., 2005; Kotecki et al., manuscript in preparation). Ubiquitous overexpression of *Gclm* was shown to have less of an effect on lifespan than *Gclc*, although the increase in longevity was significant (Orr et al., 2005). In the experiments described in this chapter, the effect of *DmGS* overexpression in *Drosophila* was studied to establish its role in determining the GSH content in flies, as well as its possible effect on oxidative stress resistance. *DmGS* was overexpressed along with the catalytic subunit of GCL to determine the relative effects of the two major enzymes involved in GSH production, with the hypothesis that increased GSH levels would increase oxidative stress resistance.

6.2 Results

6.2.1 Effects of DmGS copy number on glutathione production and resistance to oxidative stress

DmGS exists as a tandem duplication in the majority of *Drosophila melanogaster* strains, though as discussed in Chapter 3 some strains contain a single copy of the gene. In the case of strains with the GS duplication it was unclear if both copies are active and if they produce more glutathione than the single copy strains, and consequently, whether the *DmGS* duplication would confer increased oxidative stress resistance to those strains. Several strains were tested for glutathione titre and resistance to oxidative stress (Table 6.1) including the reference single and double copy strains, *w sn*^{3}, and *y sn*^{3} *v* respectively. The *w sn EP1322* parental strain, (referred to *EP1322* in this chapter), the balancer *FM6*,

and the putative *DmGS* mutant *l(1)8.3/FM6* (Chapter 3) were also tested because the *l(1)8.3* and *FM6* chromosomes carry a single copy and a double copy of *DmGS* respectively, putting that stock in the unique position of containing a total of 3 copies of *DmGS* in each female diploid cell. Because only female *l(1)8.3/FM6* flies were viable, only females were used in this study for comparative purposes.

Table 6.1: Genotypes, *DmGS* duplication status, and GSH levels of *D. melanogaster* strains tested in this study. Assays were performed on two independent samples in triplicate.

Genotype	*DmGS* Duplication Status	Number of *DmGS* copies (diploid)	GSH concentration (µmol GSH/mg tissue)
w sn	Single copy	2	36.9
y sn v	Double copy	4	37.2
w sn EP1322	Double copy	4	19.5
FM6	Double copy	4	28.9
w sn l(1)8.3/FM6	Single/Double copy	3	37.3

Total glutathione content was measured in these experiments (Chapter 2, section 2.19), and although it would have been useful to measure oxidised glutathione levels, it is well established in the literature that enzymatic, fluorimetric and colorimetric assays are not suitable to measure GSSG, not least because GSH is easily oxidised to GSSG during processing (Meister and Anderson, 1983; Jones, 2002; Rebrin et al., 2003). Ion-paring HPLC in combination with coulometric electrochemical detection has been found to be the most reliable and specific method to detect GSSH (Rebrin et al., 2003), however we did not have access to appropriate equipment.

The glutathione concentration in the strains tested appears to be unrelated to the *DmGS* copy number (Figure 6.1). *l(1)8.3/FM6* (3 *DmGS* copies) has the same concentration of GSH as both *y sn v* (4 *DmGS* copies) and *w sn* (2 *DmGS* copies). The strain with the lowest levels of *DmGS* is the parental *EP1322* genotype, which contained just over half the amount of GSH as *l(1)8.3/FM6* and the two reference strains, despite carrying the *DmGS* gene duplication. *FM6* females contained 78% of the GSH value from the two reference

strains and *l(1)8.3/FM6*.

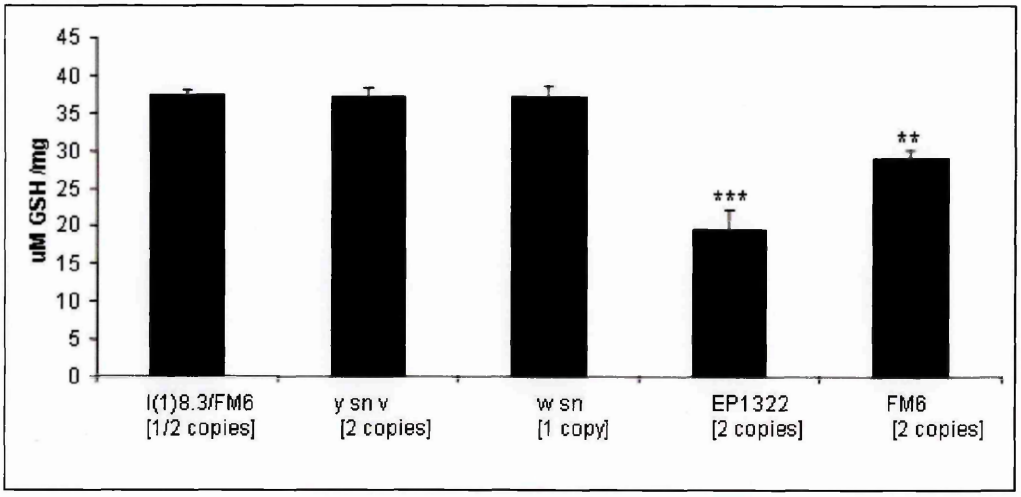


Figure 6.1: Glutathione levels in fly strains with different *DmGS* copy numbers, expressed as µmoles glutathione per mg tissue. GSH was tested in 3 day old non-virgin female flies. Results are the mean of two experiments performed in triplicate. Error bars indicate the standard error of the mean. Significant difference between *EP1322* and each of *l(1)8.3/FM6*, y sn v, and w sn; as well as between *FM6* and each of *l(1)8.3/FM6*, *y sn v*, and *w sn* was calculated using a paired Students t-test where *** = p<0.001 and ** = p<0.01.

To investigate the effect of copy number on *DmGS* expression under conditions of oxidative stress, fly strains with different copy numbers were exposed to paraquat, which generates the superoxide radical, and also to the glutathione depleting agent diethyl maleate (DEM). In these experiments, groups of three day old non-virgin flies were fed a chemical stressor, either 5mM paraquat or 6mM DEM. The number of surviving flies was counted each day, as described in Chapter 2, section 2.28. For flies of all the genotypes tested, control flies which had not been exposed to any chemical stressor had lower *DmGS* expression levels than those exposed to either paraquat or DEM (Figure 6.2), although a statistically significant induction of *DmGS* was seen only with paraquat in *l(1)8.3/FM6*.

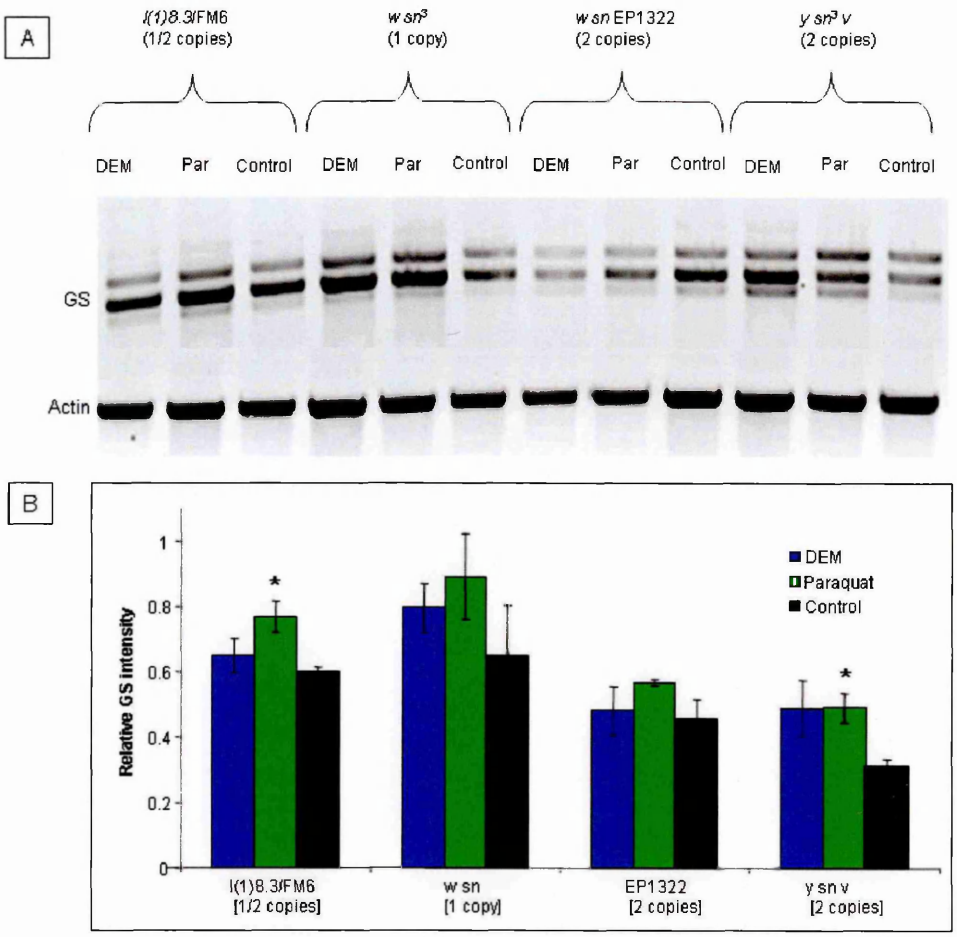


Figure 6.2: Expression of *DmGS* in flies after treatment with either DEM or paraquat. 3 day old female flies were treated with 6mM DEM or 5mM paraquat for 5 days. RNA was then extracted from the flies and RT-PCR was performed on each of the genotypes tested. The flies tested included *l(1)8.3/FM6* (has one copy of *DmGS* on one X chromosome, and two copies of *DmGS* on the second X chromosome), *w sn*^{3} (has one copy of *DmGS*), *EP1322* (has two copies of *DmGS*), *y sn*^{3} *v* (has two copies of *DmGS*). Panel A shows a sample of RT-PCR for *DmGS* and for actin5C from each of the treated fly groups. Panel B shows the intensity of pooled GS transcript bands relative to actin5C intensity. Experiments were performed in triplicate. Significant difference between treated flies and the controls was calculated using a Students t-test, where *=p<0.05. Error bars indicate the standard error of the mean.

The survival of three day old flies from strains with different *DmGS* copy numbers exposed to DEM or paraquat was calculated to determine whether *DmGS* copy number had an effect on oxidative stress resistance. *DmGS* copy number was found to have no effect on oxidative stress resistance induced by paraquat (Figure 6.3A). The least resistant and most resistant strains (*y sn*^{3} *v* and *EP1322*) both carry a duplicated *DmGS* gene, while the single copy strains (*w sn*^{3} and *l(1)8.3/FM6*) were intermediate in survival. There is also no effect of *DmGS* copy number on oxidative stress resistance in flies fed on DEM (Figure 6.3B). In

fact, strains carrying the duplicated gene appear to have decreased resistance to oxidative stress induced by DEM.

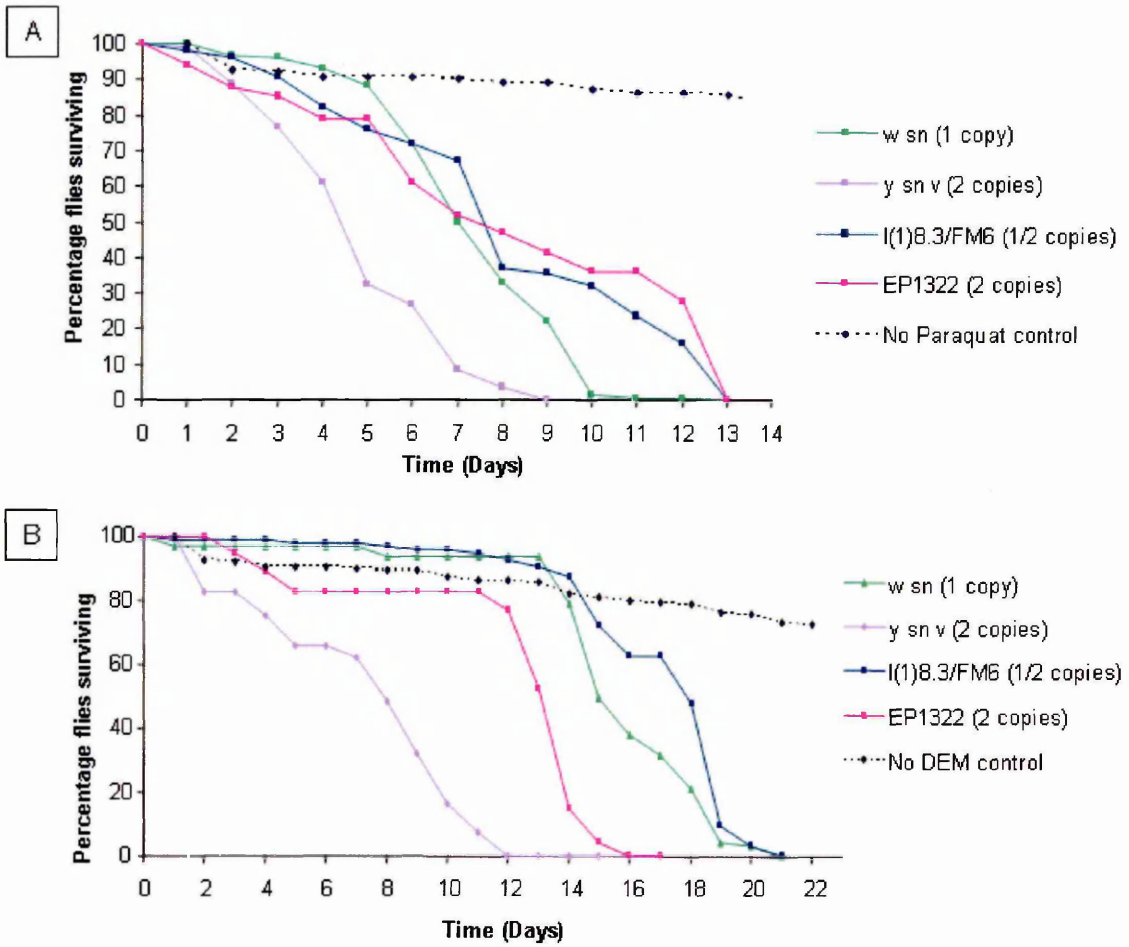


Figure 6.3: Survival curves of fly strains with different *DmGS* gene copy numbers exposed to oxidative stress. Mated 3 day old female flies were placed in a population tub, and fed on a 10% sucrose solution containing the appropriate chemical stressor (Day 0). Untreated controls were fed on 10% sucrose only, and represent the mean survival results of untreated EP1322 and l(1)8.3/FM6 flies. Panel A shows flies fed on 5mM paraquat. Panel B shows flies fed on 6mM DEM. (n=~170 for each genotype).

Lower endogenous GSH concentrations also do not correlate with decreased oxidative stress resistance in the strains tested. The double copy strain *y sn v* has the lowest resistance to both DEM and paraquat, however it has the highest amount of cellular GSH along with *l(1)8.3/FM6* (Table 6.1). In contrast, *l(1)8.3/FM6* exhibited high resistance to both chemical stressors tested.

This set of experiments is likely to be confounded by separate and unrelated genetic factors specific to each strain, and the number of strains tested is not great enough to show a pattern despite these genetic differences. To evaluate the impact of elevated *DmGS*

expression of oxidative stress resistance, the *EP1322* element was used to drive *DmGS* expression.

6.2.2 *Effect of DmGS up-regulation on glutathione production and oxidative stress survival*

Strain *EP1322* contains an EP element insertion in the 5' UTR of the distal copy of *DmGS*, upstream of the translation initiation ATG codon, as illustrated in Figure 6.4. An EP element is a 7987bp transposable P-element containing a basal promoter and 14 copies of the yeast GAL4 Upstream Activating Sequence (UAS) sequence (Kraut et al., 2001). The UAS sequences are oriented such that GAL4-dependent transcription of genomic sequences flanking the element's insertion site occurs. The GAL4 protein regulates expression of genes under the control of a UAS by binding directly to the UAS sequence (Duffy, 2002), as described in detail with reference to *Drosophila* in Section 1.3.2.3. Crossing strain *EP1322* to a strain expressing GAL4 would be expected to drive *DmGS* transcription from the UAS sequence in cells expressing GAL4.

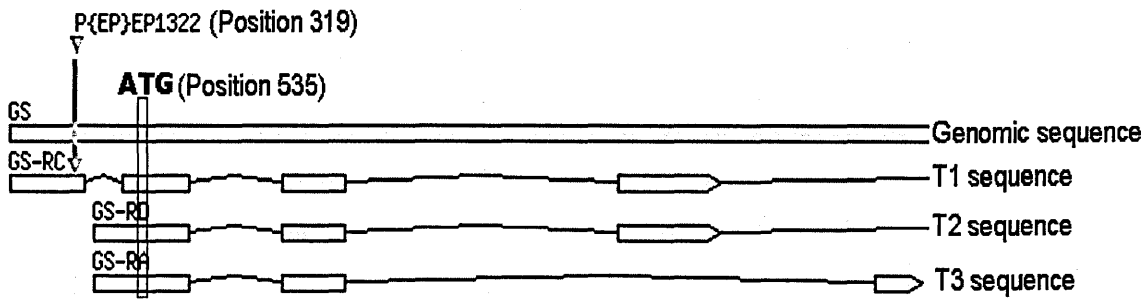


Figure 6.4: Schematic diagram of the insertion site of the EP element within the distal copy of *DmGS* in strain *EP1322*. The cartoon with the EP element insertion point marked on it is from FlyBase. The positions marked on the diagram are the nucleotide positions on the genomic sequence in Appendix B. The three transcripts shown in the diagram are those associated with the distal copy of *DmGS* (CG6835). The position of the EP element in strain *EP1322* has been published (Liao et al., 2000).

RT-PCR on flies driven with a series of GAL4 drivers

Six different strains carrying P elements which expressed GAL4 in tissue-specific patterns were used to drive expression from the *EP1322* element, in an attempt to measure increased *DmGS* transcription initiated from the *EP1322* UAS sequence. The six GAL4 driver strains chosen had distinct spatial expression patterns, or alternate induction systems, as listed in Table 6.2

Table 6.2: GAL4 drivers crossed to strain *EP1322* to measure possible changes in *DmGS* expression levels.

Stock name	Genotype	Expression pattern
Elav_GS-GAL4	P[ELAV-GeneSwitch] (3^{rd} chromosome)	Nervous system (RU486-dependent)
Actin(2)-GAL4	y^1 w^*; P[w^{+mC}=Act5C-GAL4]25FO1/CyO, y^+ (2^{nd} chromosome)	Ubiquitous
Actin(3)-GAL4	y^1 w^*; P[w^{+mC}=Act5C-GAL4]17bFO1/TM6B, Tb^1 (3^{rd} chromosome)	Ubiquitous
D42-GAL4	w^-; P[w^+, GAL4]-D42 (3^{rd} chromosome)	Adult motorneuron and salivary gland
Tubulin-GAL4	y^1 w^*; P[w^{+mC}=tubP-GAL4]LL7/TM3, Sb^1 (3^{rd} chromosome)	Ubiquitous

Males from each driver line were crossed to virgin females from the *EP1322* stock. The progeny were then tested for *DmGS* expression levels using RT-PCR. In the case of the Elav-GS GAL4 driver, which utilises the GeneSwitch system , binding of GAL4 to UAS is dependent on the presence of the activator molecule RU486 (mifepristone). Progeny from the Elav-GS driver crossed to *EP1322* were either fed on RU486 from larval or adult stage, or not at all. The protocol for administering RU486 to flies and the mechanism of the GeneSwitch system is described in section 2.29.

Of the five GAL4 drivers tested, the ubiquitously expressing Tubulin-GAL4 driver in combination with *EP1322* resulted in the highest *DmGS* expression level in both female and male progeny relative to the parental *EP1322* strain (Figure 6.5). The increase in

DmGS expression levels in the Tubulin-GAL4>*EP1322* progeny is particularly evident in female flies (~4 fold increase), although *DmGS* expression in male progeny is approximately double that of the parental *EP1322* strain. *DmGS* expression in the progeny of Elav-GS>*EP1322* and D42-GAL4>*EP1322* did not show any elevation compared to the parental strain. This could be a result of the drivers' genetic background, or it may be because any increase in *DmGS* expression in the neural tissues alone would not be sufficient to be measured by RT-PCR of whole body samples.

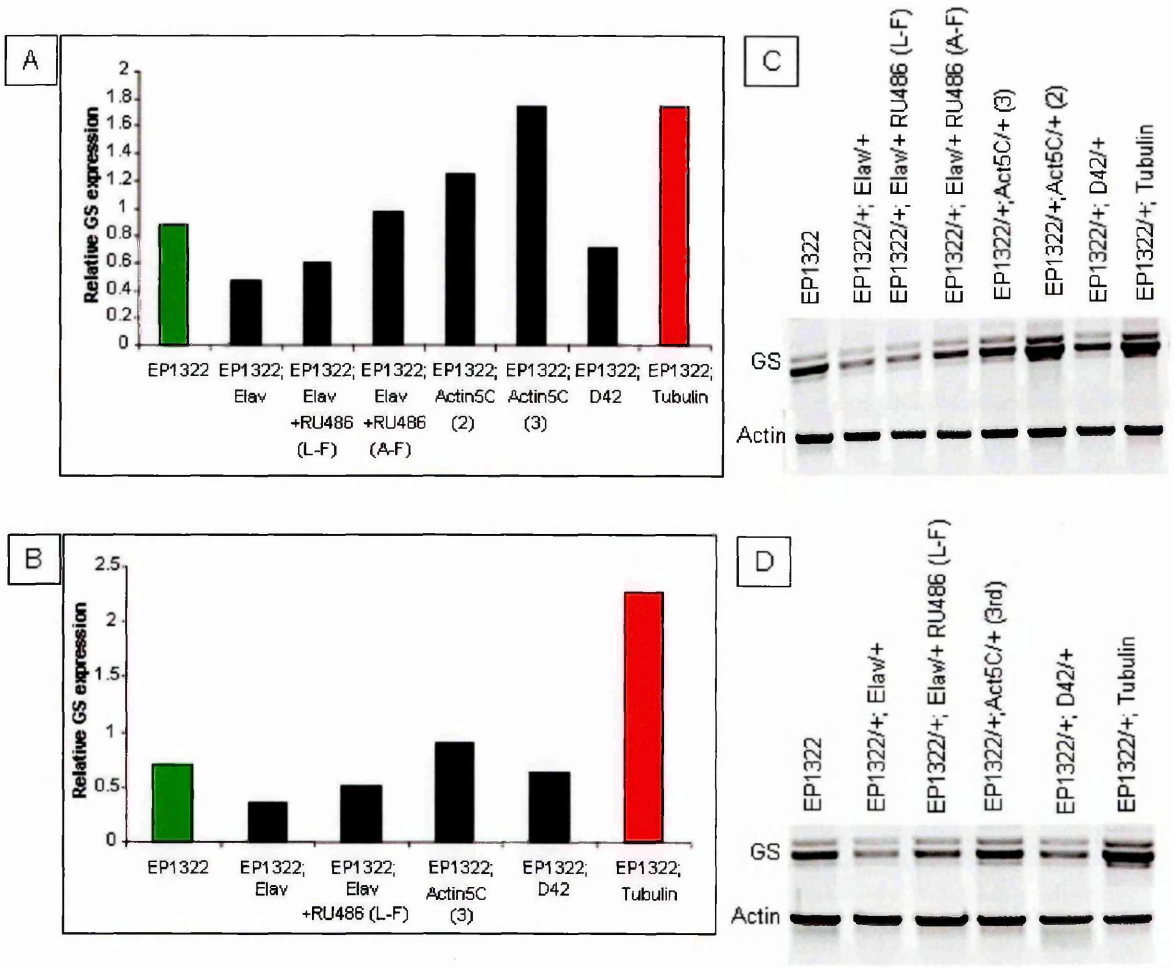


Figure 6.5: RT-PCR data showing *DmGS* expression relative to actin5C. *EP1322* flies were crossed to a series of GAL4 drivers, and the GS expression in the driven progeny was compared to the undriven control. Panels A and B show the relative expression of GS in male and female flies respectively. The Elav flies fed on RU486 were either given the RU486 from the larval stage through to adult (L-F) or in the adult stage only (A-F). The driver which resulted in the highest GS expression levels in male as well as female flies (Tubulin-GAL4) is indicated by a red bar, while the undriven *EP1322* stock is marked with a green bar. Panels C and D show the original agarose gel images of GS and actin5C RT-PCR samples of male and female flies from which the data in panels A and B were obtained. This data is from a single experiment.

Tubulin-GAL4 driven *EP1322* flies were examined more closely, as there appeared to be upregulation of *DmGS* in both female and male driven flies.

Tubulin-GAL4 induced overexpression of *DmGS* and/or *Gclc*

A *UAS-Gclc* transgene created previously (Pushpa Kansagra, Open University, pers. comm.) was included in experiments aimed at evaluating the impact of GAL4 driven expression of GS in *EP1322*, using the Tubulin-GAL4 driver. The UAS-*Gclc* transgene is fused to a UAS sequence which permits GAL4 initiation of transcription. *Gclc* was over-expressed without the modifier component, as ubiquitous high level simultaneous over-expression of *Gclc* and *Gclm* using the Tubulin-GAL4 driver has been found to be lethal (Orr et al., 2005; Kotecki et al., manuscript in preparation). The *Gclc* transgenic stock is transformed with a pUAST construct bearing a complete *Drosophila Gclc* cDNA: the resulting transformant stocks express *Gclc* in response to GAL4 driver elements (Fraser et al., 2002). A stock carrying both the *EP1322* P-element and the pUAST-*Gclc* transgene was constructed (Figure 6.6). The markers used to create the stock included *CyO* and *Sp*. Both *EP1322* and UAS-*Gclc* elements are marked with a mini-*white* marker, allowing identification of flies bearing both elements on the basis of darker eye pigmentation compared to single element flies. To create a stock homozygous for pUAST-*Gclc*, the males in step 6 were crossed to female siblings carrying pUAST-*Gclc* and *FM6*. The stock was eventually maintained as *EP1322*; pUAST-*Gclc*, although experiments were performed on flies with the *EP1322*/FM6; pUAST-*Gclc* genotype.

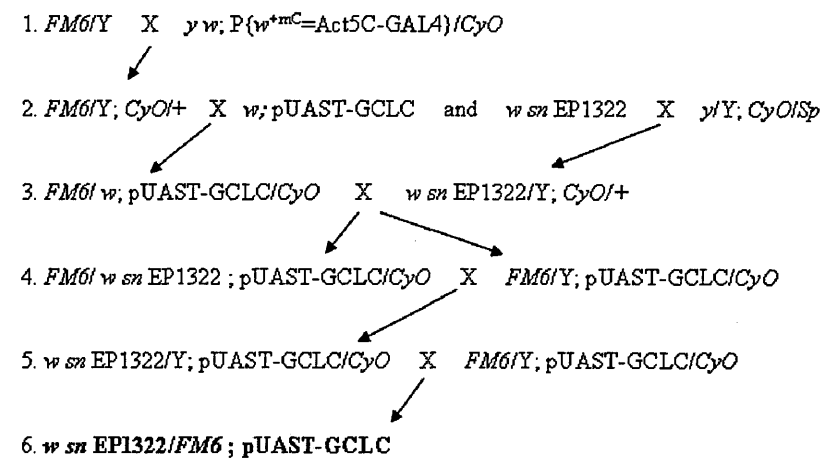


Figure 6.6: Crossing scheme to introduce a construct containing a UAS-*Gclc* transgene into the *EP1322* stock. In step 5 heterozygous and homozygous pUAST-*Gclc* carrying flies were distinguished by a darker eye colour in the heterzygotes as a result of the mini*white* gene carried by pUAST.

The crossing scheme to create Tubulin-GAL4 driven *EP1322* flies with or without the presence of a transgenic copy of UAS-*Gclc* is illustrated below in Figure 6.7. Darker eyed flies from the *EP1322* cross were collected as there is a w^+ marker in both the EP1322 and GAL4 driver elements, as well as in the pUAST-*Gclc* construct, resulting in a darker eyed phenotype in flies carrying all three elements.

♂ Tubulin-GAL*4/TM3, Sb* X ♀ *w sn*^{3} *EP1322*
or ♀ *w sn*^{3} *EP1322/FM6;*
or ♀ pUAST-GCLC
↓

w sn^3 *EP1322/+*; Tubulin-GAL*4/+*
or *w sn*^{3} *EP1322/+*; Tubulin-GAL*4*/pUAST-GCLC
or pUAST-GCLC/Tubulin-GAL*4*

Figure 6.7: Crossing scheme to introduce a Tubulin GAL-4 driver to three stocks containing combinations of *EP1322* and UAS-*Gclc*.

Levels of *DmGS* transcription were measured in the three driven stocks described in Figure 6.7, as well as in the tubulin-GAL4 driver stock and the parental pUAST-*Gclc* and *EP1322* stocks. The genotypes of the flies used in all following experiments are given in Table 6.3.

Table 6.3: Genotypes of flies used in the oxidative stress experiments that follow

Fly name	Genotype
Tubulin-Gal4 Driver (TubGAL4)	*y w*; P[*w* =tubP-GAL4]LL7/TM3, *Sb*^{1}
EP1322	*w sn EP1322*
UAS-*Gclc*	*w*^{1118}; pUAST-*Gclc*
EP1322; Gclc	*w sn EP1322/FM6;* pUAST-*Gclc*
Driven *EP1322*	*w sn EP1322/+;* TubGAL4/+
Driven *Gclc*	pUAST-*Gclc*/TubGAL4
Driven *EP1322;Gclc*	*w sn EP1322/+;* pUAST-*Gclc*/TubGAL4

RNA was extracted from 3day old male and female flies, and cDNA was created and amplified in PCR reactions as described in Chapter 2, sections 2. 4.2 and 2.16. *DmGS* expression was increased in male and female *EP1322* ; tubulin-GAL4/+ flies compared with the undriven parental strain *EP1322* (Figure 6.8). The increase in *DmGS* expression

was greater in *EP1322*;tubulin-GAL4/+ compared with *EP1322*;tubulin-GAL4/UAS-*Gclc*. In all cases *DmGS* was more highly expressed in males than in females, although the reason for this is not known at this stage. In order to rule out the possibility that the presence of the tubulin-GAL4 driver alone could raise *DmGS* expression, expression of *DmGS* was tested in tubulin-GAL4/UAS-*Gclc* flies which did not carry the *EP1322* element. There was no increase in *DmGS* expression in these stocks, confirming that the increase in *DmGS* expression seen in driven *EP1322* stocks is due to the GAL4-induced activation of *DmGS* expression mediated by the EP element *EP1322*.

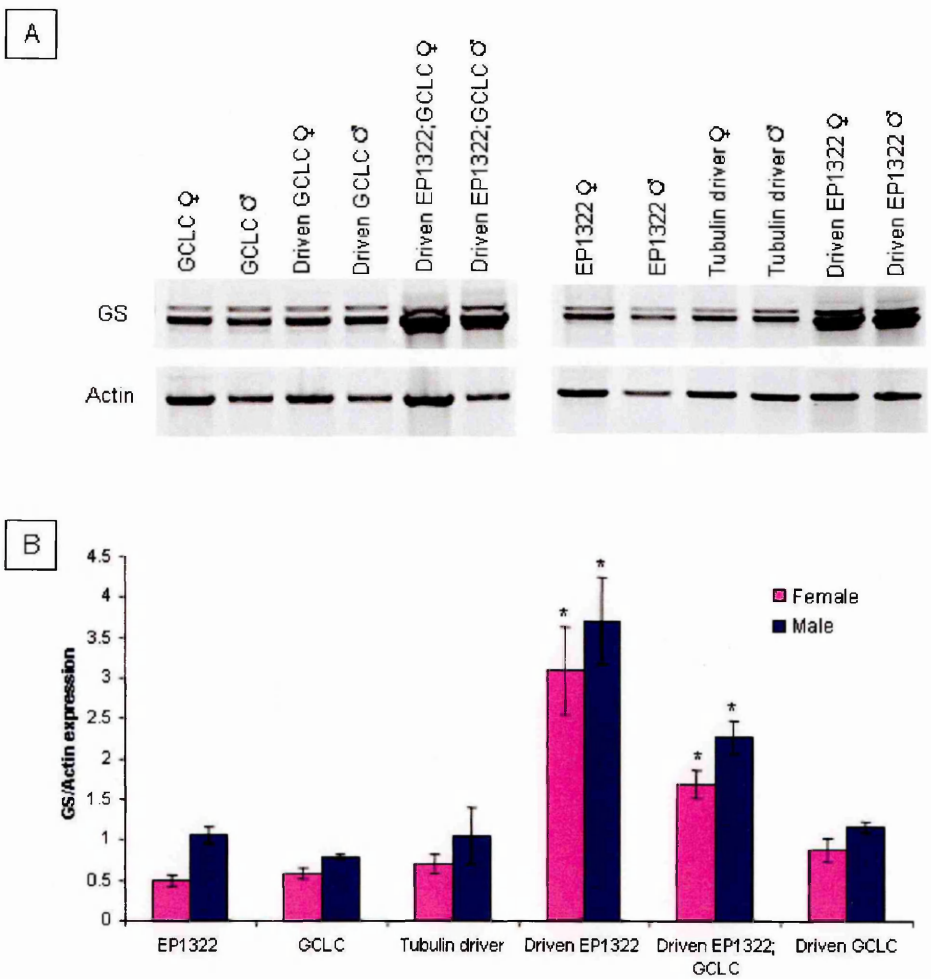


Figure 6.8: RT-PCR data showing GS expression relative to actin5C. *EP1322* flies with and without a *Gclc* transgene were crossed to Tubulin-GAL4 driver, and the GS expression in the driven progeny was compared to the undriven control. Panels A shows the GS and actin5C RT-PCR samples from male and female driven and control *EP1322* and *Gclc* flies. Panel B shows the expression of GS relative to actin5C in male and female driven and control *EP1322* and *Gclc* fly strains. Statistically significant difference between the GS levels in both male and female driven EP1322 flies (Driven EP1322; GCLC and Driven GCLC) and their non-driven EP1322 counterparts (EP1322) is indicated with an asterisk (Students t-test; *p<0.05).

The level of *Gclc* expression was also measured in the genotypes discussed above. The *EP1322*; UAS-*Gclc* genotypes crossed to the tubulin driver had at least a 3-fold increase in *Gclc* expression compared to the undriven parental genotypes (Figure 6.9). As is the case with *DmGS* expression, driven *EP1322*; UAS-*Gclc* males exhibited an unexplained higher level of *Gclc* transcription compared to their female siblings.

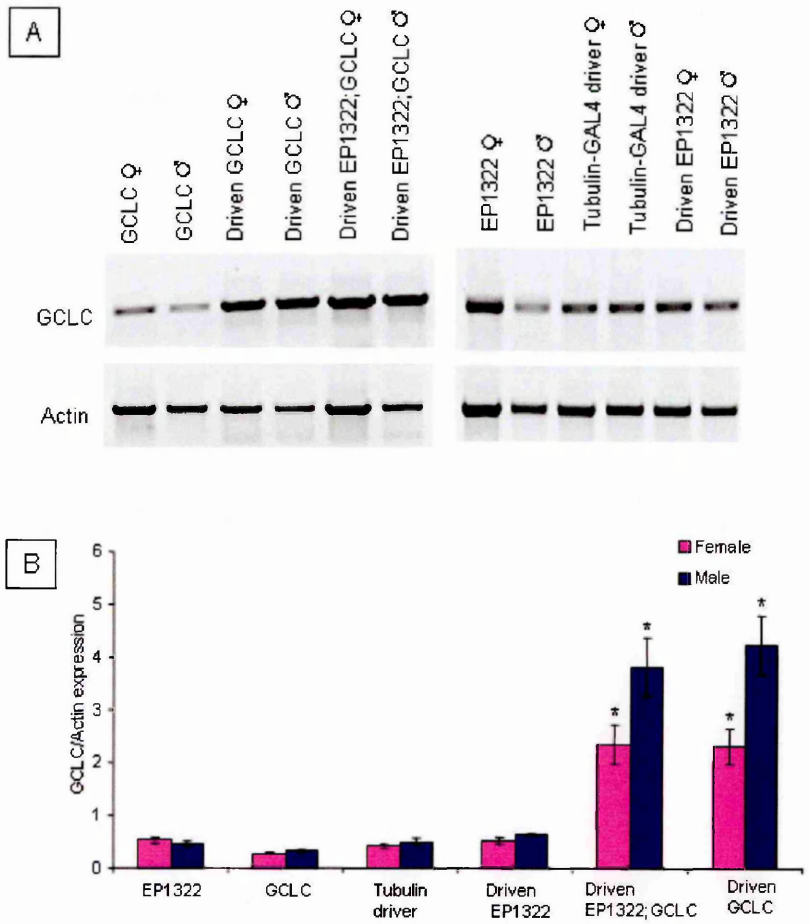


Figure 6.9: RT-PCR data showing *Gclc* expression relative to actin5C. *EP1322* flies with and without a *Gclc* transgene were crossed to Tubulin-GAL4 driver, and the *Gclc* expression in the driven progeny was compared to the undriven control. Panel A shows *Gclc* and actin5C RT-PCR samples from driven and control male and female flies. Panel B shows the expression of GCLC relative to actin5C in male and female driven and control *EP1322* and *Gclc* flies. Statistically significant difference between the *Gclc* levels in both male and female driven *Gclc* flies (Driven EP1322; GCLC and Driven GCLC) and their non-driven *Gclc* counterparts (GCLC) is indicated with an asterisk (Students t-test; *p<0.05).

Glutathione levels in flies overexpressing *DmGS* and *Gclc*.

After confirming that *DmGS* expression was increased in *EP1322* ; tubulin-GAL4/+, and that *Gclc* expression was increased in UAS-*Gclc*/Tubulin-GAL4 flies, the glutathione content of these flies was measured as described (Chapter 2, section 2.19) in order to ascertain whether increased expression of the genes encoding components of the glutathione synthetic pathway was reflected by increased GSH synthesis.

All the female flies exhibiting increased transcription of *DmGS* and/or *Gclc* had greater concentrations of GSH than their non-driven counterparts (Figure 6.10A). The difference in glutathione levels among the driven flies and their parental stocks is shown to be statistically highly significant (Figure 6.10B-D). The highest GSH level was seen in *EP1322;* UAS-*Gclc*/Tubulin-GAL4 flies, suggesting a cumulative effect of *DmGS* and *Gclc* on GSH levels (Table 6.4). The GSH content in these flies was significantly different to that in both the *EP1322*; Tubulin-GAL4/+ and UAS-*Gclc*/Tubulin-GAL4 flies (Figure 6.10E). *EP1322*; Tubulin-GAL4/+ flies had the lowest GSH levels of all three driven flies, suggesting GS has less impact on physiological GSH concentrations than does GCLC.

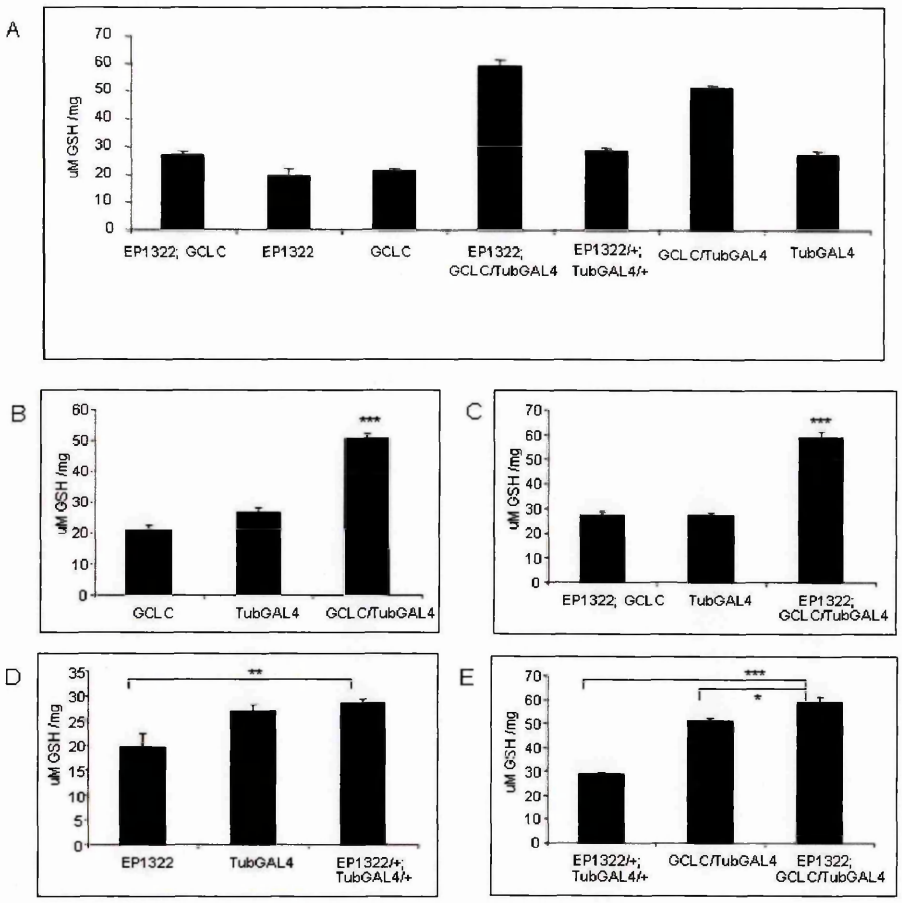


Figure 6.10: Glutathione concentrations in female flies with increased levels of GS or *Gclc* or GS+*Gclc*. Glutathione was measured in 3 day old mated flies. Glutathione levels are expressed as µmoles glutathione per mg of fly tissue, thus removing the bias of body size differences between genotypes and sexes. Panel A=all transgenic fly strains, alongside the control flies and the tubulin-GAL4 driver. Panel B=Levels of glutathione in driven and control female flies carrying the *Gclc* transgene. Panel C=Levels of glutathione in control and driven *EP1322*; *Gclc* flies which exhibited increased GS and *Gclc* expression levels. Panel D= Levels of glutathione control and driven *EP1322* flies which exhibited increased GS expression levels. Panel E=Compares the amount of glutathione present in flies overexpressing GS, *Gclc*, and GS+*Gclc*. The asterisks indicate a statistically significant difference between the samples using a Students t-test, where *=p<0.05, **=p<0.01, ***=p<0.001

Table 6.4: Glutathione levels in *Gclc* and *DmGS* overexpressing flies

Fly strain	Glutathione (µmoles/mg tissue)	Sex	Increase in GSH in driven flies vs non-driven
pUAST-*Gclc*	21.08	Female	
pUAST-*Gclc*	20.75	Male	
EP1322	19.48	Female	
EP1322	19.76	Male	
EP1322; pUAST-Gclc	27.24	Female	
EP1322; pUAST-Gclc	30.14	Male	
Tubulin-GAL4 driver	27.09	Female	
Tubulin-GAL4 driver	24.72	Male	
Driven *EP1322*	28.47	Female	49.5%
Driven *EP1322*	31.01	Male	63.2%
Driven pUAST-*Gclc*	51.03	Female	142.9%
Driven pUAST-*Gclc*	70.22	Male	243.2%
Driven *EP1322*; pUAST-*Gclc*	58.88	Female	118.5%
Driven *EP1322*; pUAST-*Gclc*	68.78	Male	128.2%

In a similar pattern to the female flies, males with increased *DmGS* and/or *Gclc* expression had greater concentrations of GSH than their non-driven counterparts (Figure 6.11A and Table 6.4). The difference in glutathione levels among the driven flies and their parental stocks was statistically highly significant (Figure 6.11B-D). In contrast to the female siblings, the highest GSH levels were seen in flies carrying a transgenic copy of *Gclc*, where the presence of the EP element *EP1322* did not appear to have an effect on the amount of GSH present. The GSH content in the *EP1322*/+; UAS-*Gclc*/Tubulin-GAL4 flies was not significantly different to the UAS-*Gclc*/Tubulin-GAL4 flies, however both these flies had a statistically significant higher level of GSH when compared to undriven parental genotypes (Figure 6.11E). When comparing the GSH levels of male and female siblings of the genotypes in which expression of *DmGS* and *Gclc* is driven by Tubulin-GAL4, male flies always had higher levels of GSH (Table 6.4). This is consistent with the higher *DmGS* and *Gclc* expression levels in male flies.

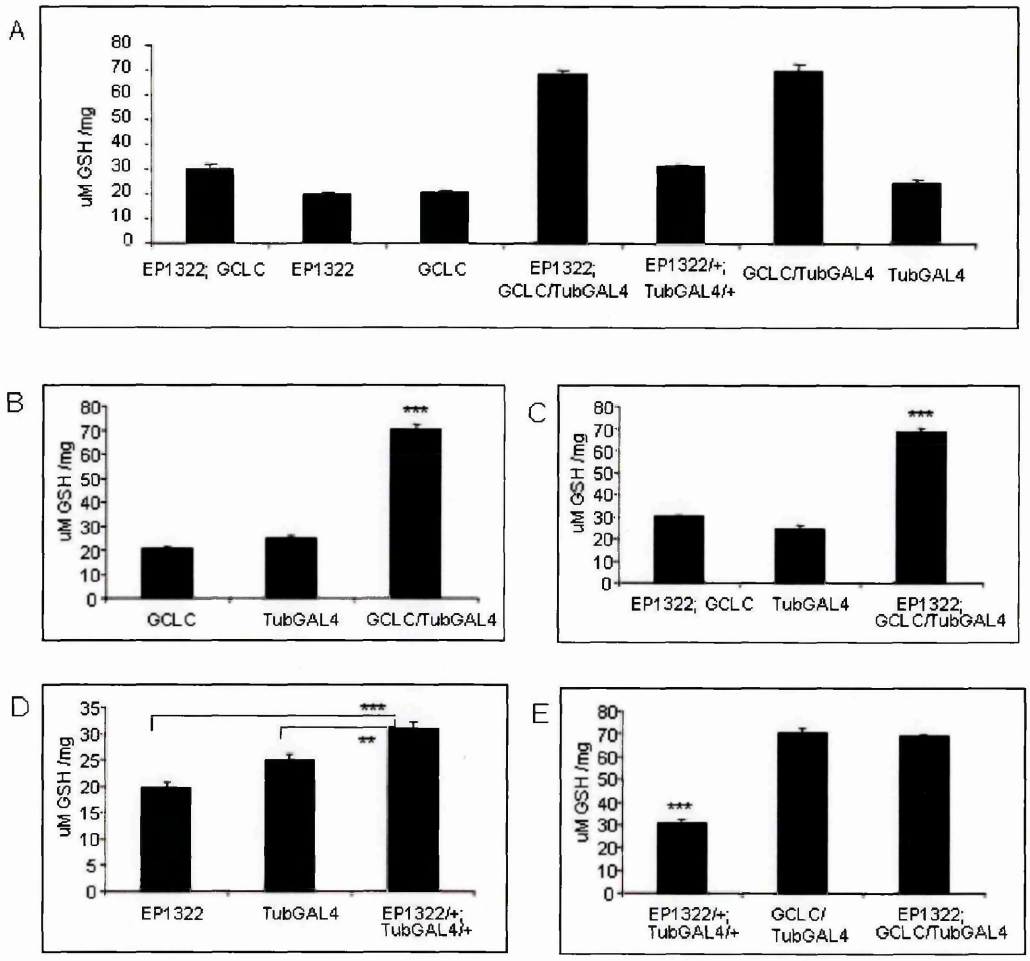


Figure 6.11: Glutathione concentrations in male flies with increased levels of *DmGS* or *Gclc* or *DmGS+Gclc* expression. Panel A=all transgenic fly strains, alongside the control flies and the tubulin driver. Panel B=Levels of glutathione in driven and control female flies carrying the *Gclc* transgene. Panel C=Levels of glutathione control and driven *EP1322*; *Gclc* flies which exhibited increased *DmGS* and *Gclc* expression levels. Panel D= Levels of glutathione control and driven *EP1322* flies which exhibited increased *DmGS* expression levels. Panel E=Compares the amount of glutathione present in flies overexpressing *DmGS*, *Gclc*, and *DmGS+Gclc*. The asterisks indicate a statistically significant difference between the samples using a Student's t-test, where **=p<0.01, ***=p<0.001

Survival of flies exposed to oxidative stress

After confirming that driving *DmGS* expression from the *EP1322* UAS sequence using a tubulin-GAL4 driver resulted in raised *DmGS* mRNA levels, and that this increase in *DmGS* expression led to an increase in GSH levels in female flies, the driven flies were tested for oxidative stress resistance. Flies overexpressing *Gclc*, as well as flies overexpressing both *DmGS* and *Gclc* (*GS+Gclc*) were also tested in this survival assay to determine whether *DmGS* could have any effect alone, or whether both enzymes involved in glutathione biosynthesis would need to be present for an effect to be observed. The chemicals used to induce oxidative stress were DEM and paraquat, fed to the flies in a

sucrose solution as described in (Chapter 2, section 2.28). The tests were performed on female and male siblings, and survival experiments were repeated at least twice in tubs containing a minimum of 75 individuals. The data from all repeated experiments was pooled to allow for greater power of statistical analysis.

Survival of flies fed on DEM

Flies were fed a 6mM DEM solution made up in 10% sucrose. DEM is a glutathione depleting compound that binds directly to GSH, and also indirectly affects glutathione content by interfering with glutathione S-transferase (Parkinson, 2001; Yang et al., 2004). DEM therefore induces oxidative stress due to glutathione depletion. It is also an interesting stressor to use in this project as it allows the specific study of a low level of glutathione, and the effect of increasing the glutathione biosynthetic enzymes in this context.

The positive consequence of raising *DmGS* expression levels under conditions of DEM-induced oxidative stress is noticeable in female flies (Figure 6.12B&C), whereas in males the induction of *DmGS* expression appears to have no effect on survival (Figure 6.13B&C). Female flies overexpressing *DmGS*, with or without *Gclc* overexpressed have an increased resistance to oxidative stress. Both male and female flies overexpressing *Gclc* do not fare any better than the non-driven control flies (Figure 6.12A and Figure 6.13A). When comparing the survival curves of all driven flies, with either *DmGS* and/or *Gclc* over-expressed, there is no difference in the survival curves for female flies (Figure 6.12D). Conversely, the male siblings over-expressing *Gclc* only survive almost twice as long as those overexpressing a combination of *DmGS* and *Gclc* (driven *EP1322*/+; UAS-*Gclc*/+), or *DmGS* only (driven *EP1322*/+) (Figure 6.13D), but this extension in survival is problematic because the non-driven *Gclc* control flies also have an extended survival curve (Figure 6.13A). This suggests that there is an unidentified genetic factor in the UAS-*Gclc* stock conferring resistance to DEM-induced stress.

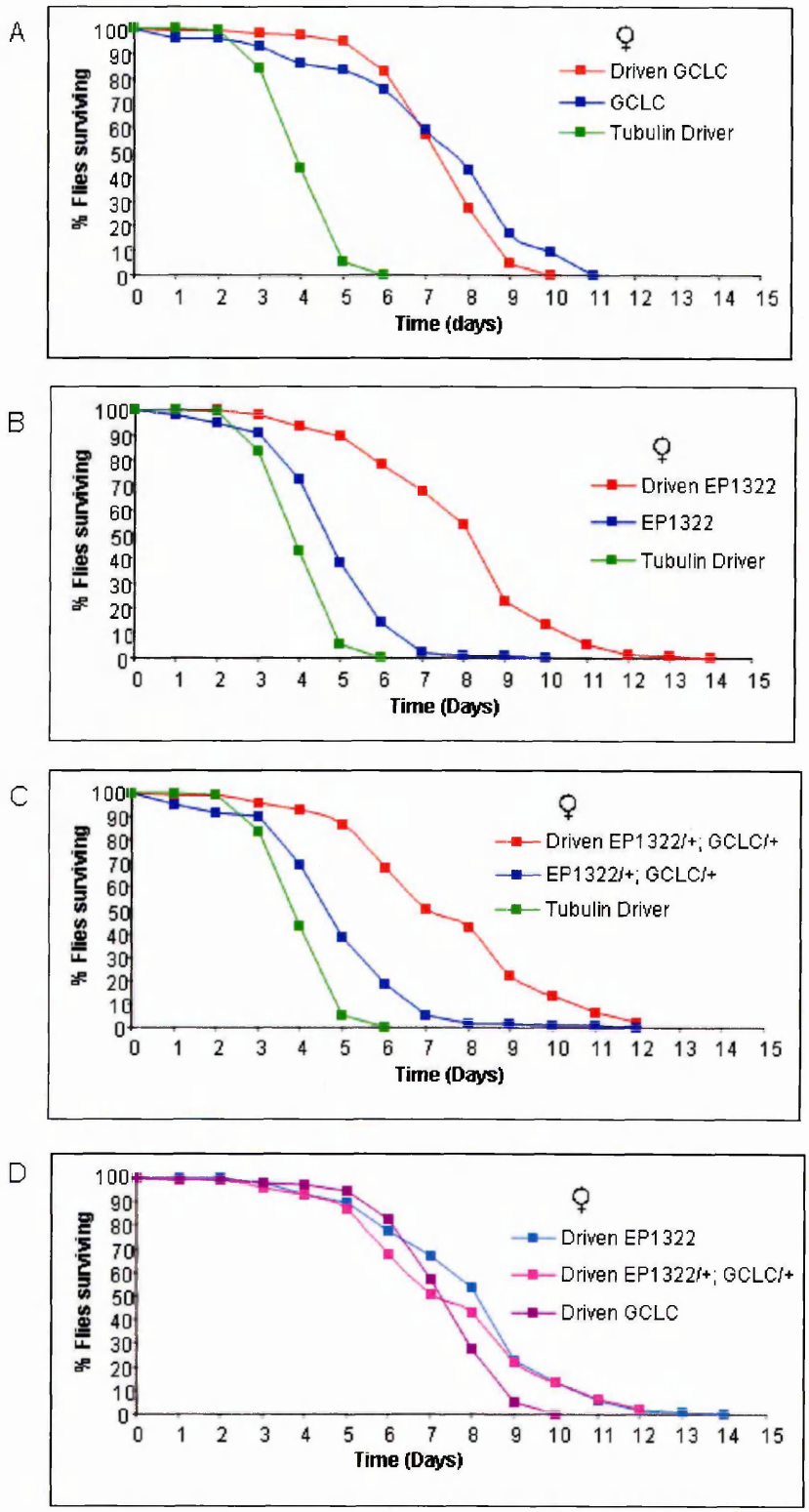


Figure 6.12: Comparison of survival curves of female flies fed on 6mM DEM. Panel A compares the survival of flies overexpressing *Gclc* compared to control flies. Panel B compares the survival of flies overexpressing *DmGS*, achieved by driving the P-element *EP1322*, compared to control flies. Panel C compares the survival of flies over-expressing both *DmGS* and *Gclc* compared to the controls. Panel D compares the survival curves of flies with different combinations of *DmGS* and *Gclc* over-expression.

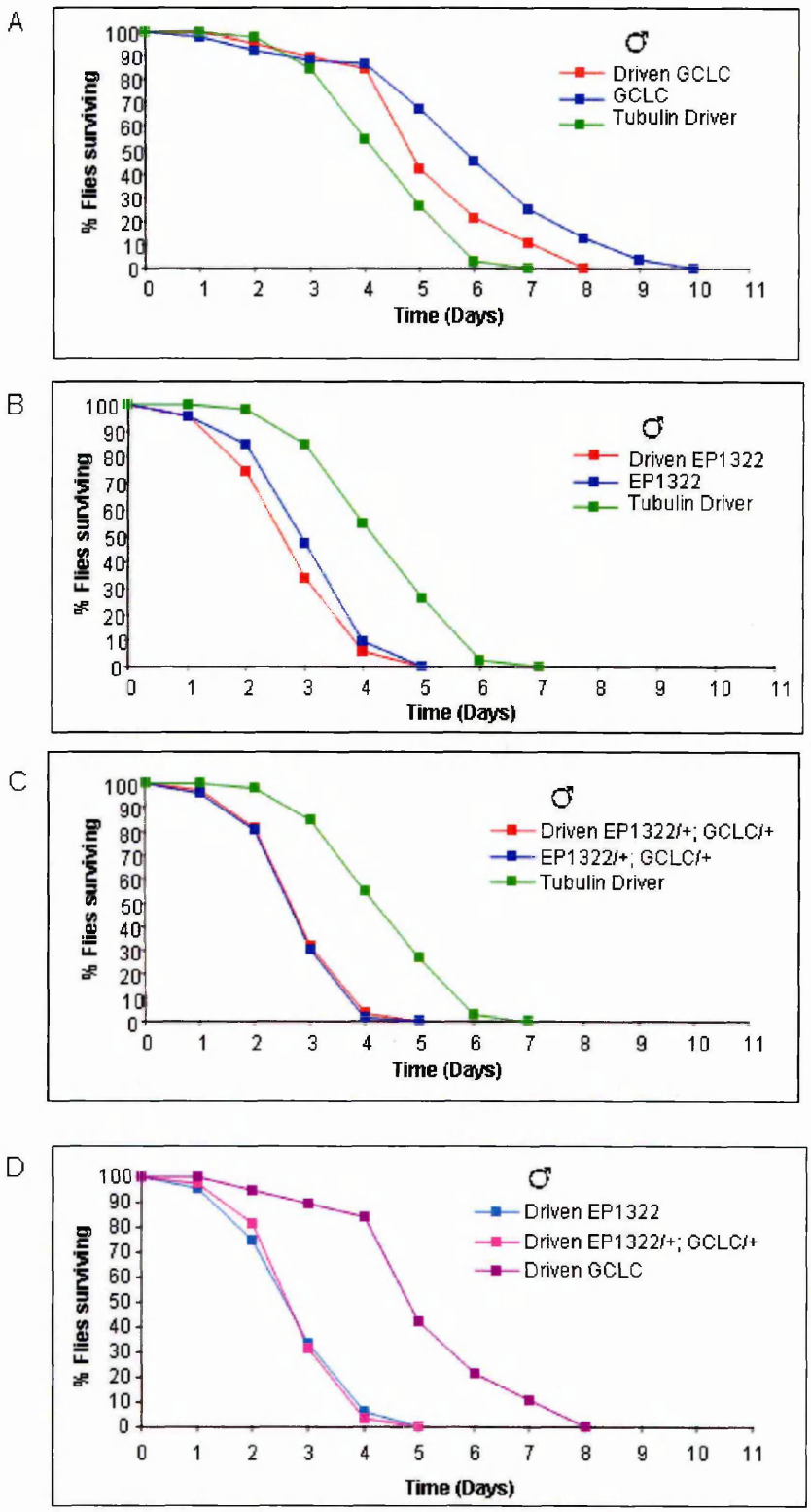


Figure 6.13: Comparison of survival curves of male flies fed on 6mM DEM. Panel A compares the survival of flies overexpressing *Gclc* compared to control flies. Panel B compares the survival of flies overexpressing *DmGS*, achieved by driving the P-element *EP1322*, compared to control flies. Panel C compares the survival of flies overexpressing both *DmGS* and *Gclc* compared to the controls. Panel D compares the survival curves of flies with different combinations of *DmGS* and *Gclc* overexpression. The full genotypes of the flies used are: Driven EP1322 = w sn EP1322/+; TubGAL4/+. Driven EP1322/+; GCLC/+ = EP1322/+; GCLC/TubGAL4. Driven GCLC = w; GCLC/TubGAL4. Driver = *y w;* TubGAL4*/TM3, Sb.*

Table 6.5: Median and maximum lifespan values of driven and non-driven flies overexpressing *DmGS* and/or *Gclc* when exposed to DEM. Values shown are±SEM.

Fly strain	Sex	Median lifespan	Maximal lifespan
EP1322	♂	2.92 ± 0.11	5.26 ± 0.10
Driven *EP1322*	♂	2.59 ± 0.06	4.66 ± 0.08
EP1322;UAS-Gclc	♂	2.6 ± 0.08	4.68 ± 0.01
Driven *EP1322*;UAS-*Gclc*	♂	2.62 ± 0.08	4.72 ± 0.07
UAS-*Gclc*	♂	5.78 ± 0.20	10.40 ± 0.16
Driven UAS-*Gclc*	♂	4.81 ± 0.27	8.66 ± 0.01
Tubulin Driver	♂	4.16 ± 0.13	7.49 ± 0.06
EP1322	♀	4.64 ± 0.12	8.35 ± 0.07
Driven *EP1322*	♀	8.11 ± 0.16	14.60 ± 0.26
EP1322;UAS-Gclc	♀	4.62 ± 0.14	8.32 ± 0.13
Driven *EP1322*;UAS-*Gclc*	♀	7.01 ± 0.27	12.62 ± 0.26
UAS-*Gclc*	♀	7.52 ± 0.28	13.54 ± 0.01
Driven UAS-*Gclc*	♀	7.22 ± 0.15	13.00 ± 0.09
Tubulin Driver	♀	3.82 ± 0.12	6.88 ± 0.05

A comparison of the time point at which 50% of the population survive (the median survival time) of driven and non-driven flies is shown in Figure 6.14 and Table 6.5). The statistical test used to compare the survival curve data was the non-parametric Mann-Whitney U test, which compares the medians of two data sets (Pallant, 2005). In female flies, the driven *EP1322* and driven *EP1322*; UAS-*Gclc* flies live significantly longer than their non-driven counterparts. This effect is unlikely to be due to the genetic background in the tubulin-GAL4 driver stock because there is no significant difference between the driver stock and the non-driven *EP1322* and *EP1322*; UAS-*Gclc* control flies. In male flies, the driven flies over-expressing *DmGS* or *Gclc* never have an increased median survival (Figure 6.14B). In the case of *EP1322* and UAS-*Gclc* flies, the driven male flies have a significantly lower survival rate, which is contrary to the expectation.

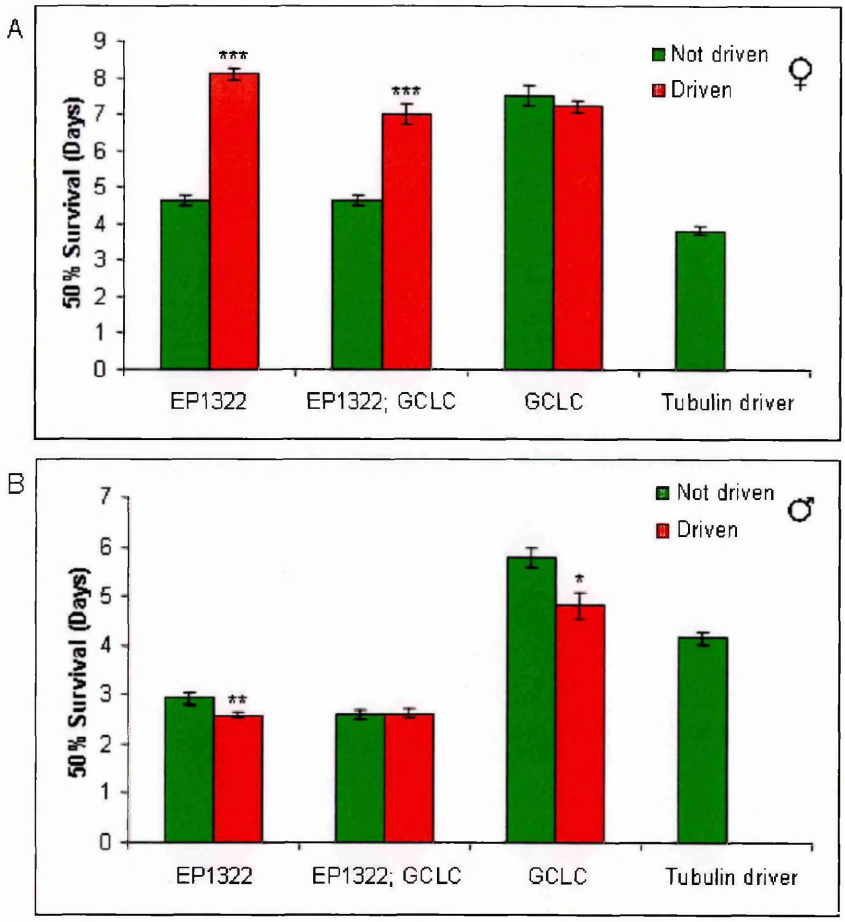


Figure 6.14: Median survival (time point at which 50% of the flies survive) in driven *EP1322* flies (red bars) over-expressing *DmGS*, *DmGS+Gclc* or *Gclc* alone, as well as control flies (green bars). All genotypes were fed on 6mM DEM. Panel A shows data for female flies, and Panel B shows data for male flies. Statistically significant differences between survival times of driven and undriven flies were calculated using the Mann-Whitney U test, where *=p<0.05, **=p<0.01 and ***=p<0.001

Survival of flies fed on Paraquat

In addition to DEM, paraquat was also used as an oxidative stressor. Paraquat is a superoxide (O_2^-) generator, and glutathione has been found to be an important cellular protector against a paraquat assault (Palmeira, 1999), making paraquat a useful source of chemically induced oxidative stress in this study. Paraquat was administered to flies in tubs in the same manner as DEM.

The survival results from paraquat-induced oxidative stress were similar to those from DEM-induced stress with respect to the sex differences. Female flies over-expressing *DmGS* and/or *Gclc* exposed to paraquat all survived for longer periods of time than the non-driven controls (Figure 6.15). Male sibling flies which over-expressed *DmGS* and/or *Gclc* did not, however, show an increase in oxidative stress resistance (Figure 6.16). The most extensive increase in survival was in females of the tubulin-GAL4 driven *EP1322*

strain, which exhibited the longest survival time of all genotypes exposed to paraquat, as well as the greatest increase in maximum survival time compared to the non-driven counterparts (Figure 6.15D).

In a comparison of all the tubulin-GAL4 driven females, the stock overexpressing *Gclc* had the lowest resistance to oxidative stress. The male flies overexpressing *DmGS* alone or with *Gclc* had similar survival curves, while the greatest resistance to oxidative stress was observed in the tubulin-GAL4 driver. Male UAS-*Gclc* flies crossed to the tubulin-GAL4 driver were not tested for paraquat resistance because there was a high level of mortality amongst the embryos of this genotype, and the numbers of flies collected were not sufficient to give statistically robust data. It has recently been suggested that global overexpression of *Gclc* may interfere with signal transduction and protein folding in a redox dependent manner (Orr et al., 2005).

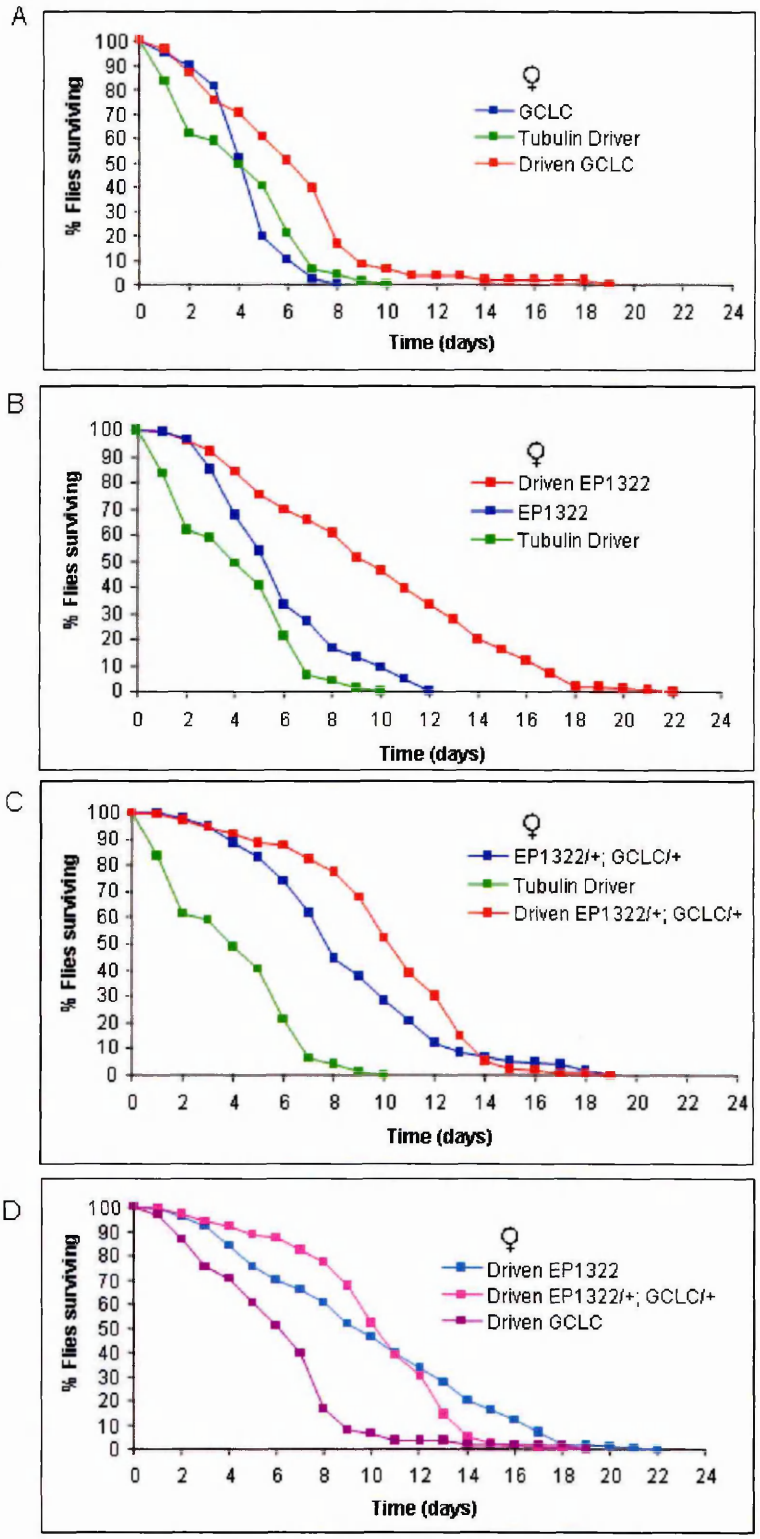


Figure 6.15: Comparison of survival curves of female flies fed on 5mM paraquat. Panel A compares the survival of flies overexpressing *Gclc* compared to control flies. Panel B compares the survival of flies overexpressing *DmGS*, achieved by driving the P-element *EP1322*, compared to control flies. Panel C compares the survival of flies over-expressing both *DmGS* and *Gclc* compared to the controls. Panel D compares the survival curves of female flies with different combinations of *DmGS* and *Gclc* over-expression.

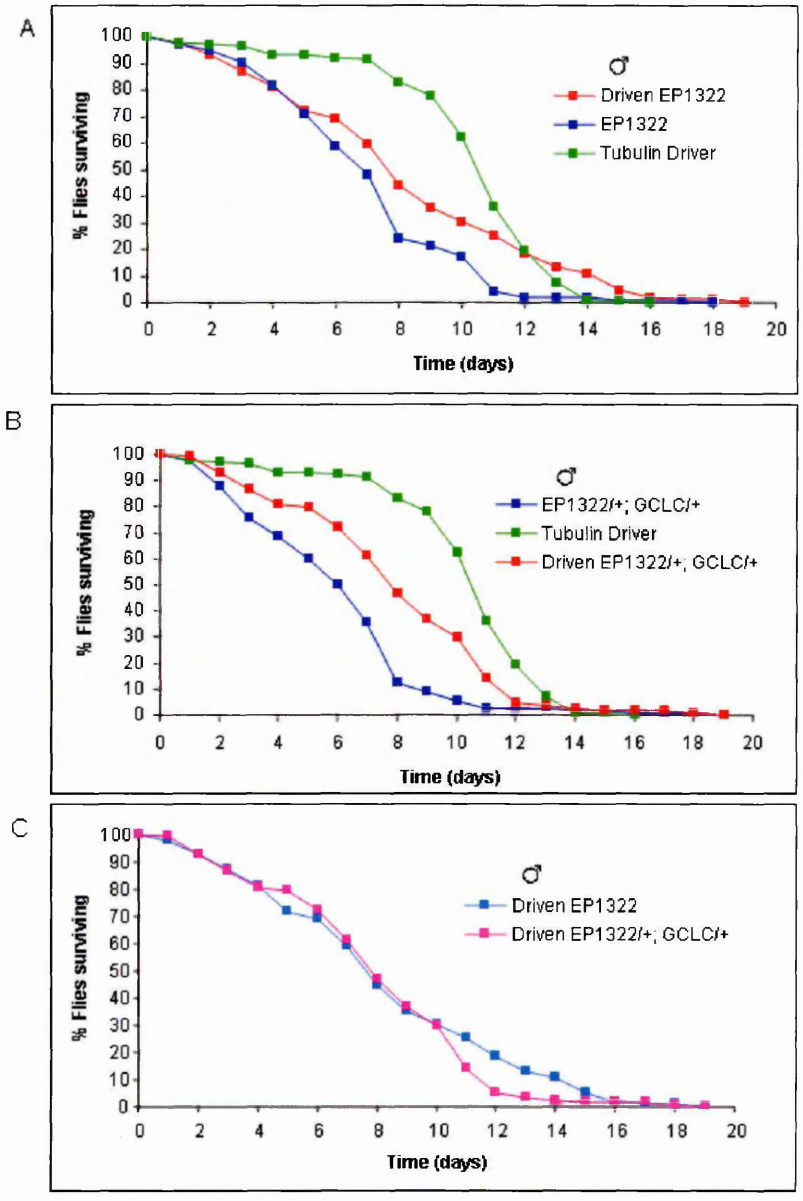


Figure 6.16: Comparison of survival curves of male flies fed on 5mM paraquat. Panel A compares the survival of flies overexpressing *DmGS*, achieved by driving the P-element *EP1322*, compared to control flies. Panel B compares the survival of flies overexpressing both *DmGS* and *Gclc* compared to the controls. Panel C compares the survival curves of male flies with either *DmGS* or a combination of *DmGS* and *Gclc* overexpression

Table 6.6: Median and maximum lifespan values of driven and non-driven flies overexpressing *DmGS* and/or *Gclc* when exposed to paraquat. Values shown are±SEM

Fly strain	Sex	Median lifespan	Maximal lifespan
EP1322	Male	6.7 ± 0.25	12.06 ± 0.19
Driven *EP1322*	Male	7.62 ± 0.33	13.71 ± 0.50
EP1322;Gclc	Male	6.01 ± 0.34	10.82 ± 0.27
Driven *EP1322;Gclc*	Male	7.76 ± 0.31	13.97 ± 0.14
Tubulin Driver	Male	10.46 ± 0.19	18.82 ± 0.23
EP1322	Female	5.19 ± 0.23	9.34 ± 0.42
Driven *EP1322*	Female	9.07 ± 0.57	16.33 ± 0.33
EP1322;Gclc	Female	7.67 ± 0.25	13.81 ± 0.76
Driven *EP1322;Gclc*	Female	10.17 ± 0.29	18.31 ± 0.20
Gclc	Female	4.05 ± 0.11	7.29 ± 0.31
Driven *Gclc*	Female	6.03 ± 0.6	10.86 ± 1.00
Tubulin Driver	Female	3.9 ± 0.51	7.02 ± 0.27

A comparison of the median survival time of driven and non-driven flies exposed to paraquat is shown in Figure 6.17 and Table 6.6. In female flies, the trend seen in the survival curves in Figure 6.15 where all driven fly strains were more resistant to paraquat than the non-driven controls is confirmed statistically using the Mann-Whitey U test. The time point where 50% of the driven population survived is significantly different to the non-driven controls (Figure 6.17A). In male flies, the driven stocks over-expressing *DmGS* with or without *Gclc* have a significantly raised survival compared to the non-driven control flies, with the flies over-expressing *DmGS* as well as *Gclc* showing a more significant difference compared to their non-driven counterparts (Figure 6.17B).

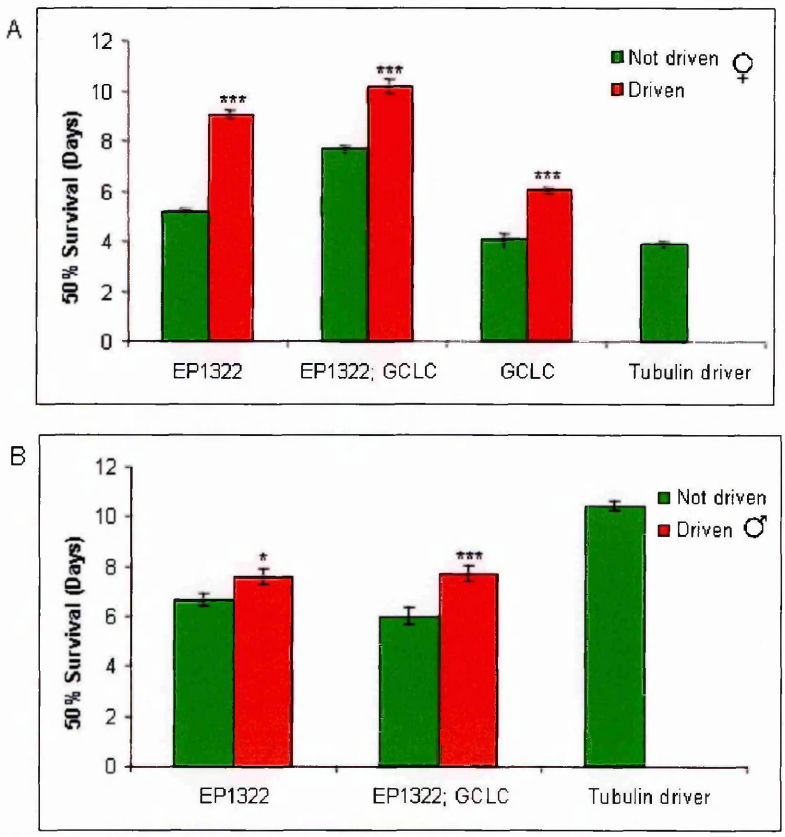


Figure 6.17: Median survival (time point at which 50% of the flies survive) in driven *EP1322* flies (red bars) over-expressing *DmGS*, *DmGS+Gclc* or *Gclc* alone, as well as control flies (green bars). All genotypes were fed on 5mM paraquat. Panel A shows data for female flies, and Panel B shows data for male flies. Statistically significant differences between survival times of driven and undriven flies were calculated using the Mann-Whitney U test, where *=p<0.05, **=p<0.01 and ***=p<0.001.

6.3 Discussion

Levels of antioxidant defences have been postulated to affect the resistance of animals to stressful environments, and in particular to environmental oxidative stress (Arking et al., 2000; Sohal et al., 2002; Dugan and Quick, 2005; Orr et al., 2005). An age-related increase in ROS production with a simultaneous decrease in glutathione is observed in a wide variety of organisms including *Drosophila* (Sohal et al., 1995; Rebrin et al., 2004; Orr et al., 2005). The mechanisms of age-related GSH decline are not understood, but may be due to increased demand from increased ROS, or from a decrease in *de novo* synthesis (Zhu et al., 2006). A recent study where the capacity for glutathione biosynthesis was enhanced by overexpressing GCL was shown to increase lifespan by up to 50%, and also increase resistance of flies to oxidative stress resistance (Orr et al., 2005). The hypothesis that increasing the levels of the antioxidant GSH can increase oxidative stress resistance was tested in this study by genetically manipulating levels of glutathione in flies and exposing them to oxidative stress. Flies over-expressing *DmGS* and/or *Gclc* were exposed to two oxidative stressors to assess the effects of increased expression of both glutathione biosynthesis genes on artificially induced oxidative stress resistance.

The chemicals used to induce oxidative stress were DEM and also paraquat. DEM is a glutathione depletor, lowering GSH by over 85% at 1mM concentrations (Ito et al., 1998; Parkinson, 2001). DEM is used to chemically induce oxidative stress because GSH depletion causes an increase in free radical generation, although the mechanism for the increase in ROS is still unclear (Back et al., 1998). Superoxide radicals are generated when paraquat is reduced by NADPH-diaphorase to a paraquat radical, which in turn reacts with dioxygen to produce superoxide radicals (Bus and Gibson, 1984; Orr and Sohal, 1993). Paraquat exposure is used as a non-chronological bioassay of ageing in *Drosophila* where lifespan and Paraquat resistance have been shown to track together in

laboratory strains (Arking et al., 1991; Ruan et al., 2002; Morrow et al., 2004; Bauer et al., 2005; Droge, 2005) as well as in wild caught flies (Baldal et al., 2006).

A preliminary study was performed looking at *DmGS* copy number and *DmGS* expression in relation to oxidative stress in a variety of non-driven *Drosophila* strains, however there was no correlation between *DmGS* copy number, *DmGS* expression and GSH content. *DmGS* copy number also had no effect on oxidative stress resistance when the flies were exposed to DEM or paraquat. The results of this experiment are probably confounded by separate and unrelated genetic factors specific to each fly strain, and prompted the use of GAL4 driven flies where the genetic background would be more comparable. This experiment did, however, confirm that *DmGS* expression was increased in all genotypes when exposed to DEM as well as paraquat, although the trend was only statistically significant in *l(1)8.3/FM6* flies. A previous study also found an increase in *DmGS* expression in 3 day old male flies exposed to paraquat compared to the control flies (McCarroll et al., 2004), while DEM has been shown to upregulate *GS* in rat hepatocytes (Huang et al., 2000). DEM and paraquat-induced oxidative stress are thought to increase GSH by a mechanism that involves the redox sensitive transcription factors AP-1 and NKκB, as well as transcription factors Nrf1 and Nrf2 (Coe et al., 2002), as discussed in detail in Chapter 1, section 1.1.4.

In order to minimize the genetic background problems encountered using unrelated stocks with different *DmGS* copy numbers, transgenic flies that had raised *DmGS* and *Gclc* expression levels when crossed to a GAL4 driver were used. The *EP1322* insertion was used to drive *DmGS* expression when driven by a tubulin-GAL4 driver, and a pUAST-*Gclc* insertion overexpressed *Gclc* when similarly combined with a tubulin-GAL4 driver. A recombinant strain carrying both the *EP1322* element and the pUAST-*Gclc* transgene, showed increased transcription of both *DmGS* and *Gclc* when crossed to the tubulin-GAL4 driver.

The choice of GAL4 driver was informed by the level of increase in *DmGS* transcription in driven *EP1322* flies. The tubulin-GAL4 driver caused the highest levels of *DmGS* expression in both male and female *EP1322* flies. The variability of target gene expression levels using different GAL4 drivers was also noted in a study overexpressing *ApoE* (Walker et al., 2006). Most recent studies in *Drosophila* that have shown increased stress resistance or lifespan after overexpressing a transgenic antioxidant have used a motorneuron GAL4 driver (Parkes et al., 1998; Phillips et al., 2000; Spencer et al., 2003; Hwangbo et al., 2004; Morrow et al., 2004; Fridell et al., 2005; Orr et al., 2005; Thiruchelvam et al., 2005). Here, a neuronal GAL4 driver was not used because firstly, the neuronal Elav-GAL4 driver (GeneSwitch) did not elevate *DmGS* transcription in combination with *EP1322*, and secondly, the examples cited above involved overexpression of genes that perform an antioxidant function as the single gene. GS has a secondary role in GSH synthesis, as it is not the rate limiting step. In this regard GS and GCLM are similar, and in a previous study, an increase in longevity was only achieved by overexpressing *Gclm* ubiquitously with a tubulin driver, while a neuronal expression of the active *Gclc* unit caused an increase in longevity (Orr et al., 2005).

When using the tubulin-GAL4 driver, *DmGS* expression levels were raised in driven *EP1322* flies, as well as driven *EP1322*; UAS-*Gclc* flies. Both male and female flies had increased *DmGS* expression, but higher levels of induction were seen in male flies. This sex bias was also noted in the driven UAS-*Gclc* flies where males had higher *Gclc* expression levels than their female siblings. The higher *DmGS* and *Gclc* expression levels also corresponded to higher GSH levels in driven male flies compared to their female counterparts.

Both male and female driven *EP1322* flies had the lowest GSH levels of all three genotypes of driven flies, suggesting GS has less impact on biological GSH concentrations than does GCLC. This was an expected result since kinetic studies performed largely on rat liver GS indicates that GS is not the rate limiting step in GSH biosynthesis (Huang et al.,

2000). GS is, however, an essential component of the GSH cycle, as shown when the simultaneous overexpression of *DmGS* and GCL in *E. coli* improved GSH production, while neither enzyme had an effect when expressed alone (Liao et al., 2006) Global over-expression of *Gclm* in adult flies resulted in a moderate increase in lifespan (Orr et al., 2005), even though this subunit of GCL is not essential for the production of GSH. The highest increase in GSH levels was seen in tubulin driven *EP1322*; UAS-*Gclc* and UAS-*Gclc* flies. Flies over-expressing *Gclc* both with and without *DmGS* overexpression exhibited a GSH increase of 118-243%, which is comparable to values obtained by *Gclc* upregulation in a previous study (Orr et al., 2005). In female flies, overexpressing *Gclc* in addition to *DmGS* had a cumulative effect on the increase in GSH, however this cumulative effect was not observed in male flies. Overexpressing *DmGS* and *Gclc* in rats has been found to increase GSH synthesis to levels comparable with *Gclc* overexpression alone, although this study was done using male rats only (Huang et al., 2000).

Cellular ROS are critical in diseases which can shorten lifespans, although their effect is typically not seen in standard unstressed laboratory conditions (Dugan and Quick, 2005). ROS would, however, be crucial to wild animals exposed to environmental toxins and stresses on a regular basis, for example through ingestion. It is therefore thought that in an experimental model of oxidative stress, an increase in glutathione could protect an organism against ROS damage. The hypothesis that raised GSH levels could protect against an environment of increased oxidative stress was tested by exposing flies overexpressing *DmGS* and/or *Gclc* to DEM or paraquat. Using oxidative stress resistance as a predictor of longevity in *Drosophila* has been well documented (Seto et al., 1990; Arking et al., 1991; Lin et al., 1998; Giannakou et al., 2004). Female flies overexpressing *DmGS* were shown to have increased DEM and paraquat resistance, with the effect most prominent in flies exposed to paraquat. All female flies with *DmGS* and/or *Gclc* expressed exhibited statistically significant increases in stress resistance on exposure to paraquat, with flies over-expressing *DmGS* as well as *Gclc* showing the highest median and

maximum survival times. The increase in both median (average lifespan of a population) and maximal (the maximum age to which an individual of a specific population can be expected to live) survivorship of female flies over-expressing *DmGS* and/or *Gclc* is an important result, as the implication is that both 'quality of life', as well as length of life could be extended through the induction of both *DmGS* and *Gclc* in flies. The idea that an increase in mean mortality could indicate an increased 'quality of life' in *Drosophila* was recently proposed by Ruan et al. (2002). It is interesting to note that none of the male flies over-expressing *DmGS* or *Gclc* exhibited any resistance to either DEM or paraquat, especially since male flies had higher *DmGS* and *Gclc* expression, as well as higher GSH levels than the females. In a study where *Gclc* was over-expressed in male flies only, Orr et al (2005) only saw an increase in the oxidative stress resistance of 40 day old flies, and not in the younger 10 day old flies. It may be the case that the positive effect of *DmGS* or *Gclc* overexpression on stress resistance is only seen in older male flies, so the effect would not be seen in the 3 day old flies used in this current study.

This sex bias has been noted in several studies on oxidative stress and longevity in fruit flies (Lints et al., 1983; Carey et al., 1995; Spencer et al., 2003; Forbes et al., 2004; Magwere et al., 2004; Valenzuela et al., 2004; Magwere et al., 2006). In some of the studies, though, it could be argued that the increase in female survival is due to infertility caused as an additional effect of a single gene mutation conferring increased oxidative stress resistance, as is the case with the *chico* mutant (Clancy et al., 2001). This sex-specific effect on longevity has also been noticed in other species including humans (Hazzard, 1986; Lio et al., 2002; Christiansen et al., 2006), mice (Jackson et al., 2002), spiders (Burger and Promislow, 2004) and medflies (Carey et al., 1995). It has been proposed that longevity factors may be sex-specific, with males and females employing different biological strategies (Christiansen et al., 2006).

In a review of 218 studies where the lifespan of female and male *Drosophila* were compared, the mean lifespan of female flies was found to be greater than that of the males

in approximately half of the cases; however, the maximal lifespan was consistently higher for females than it was for males (Lints et al., 1983). Interventions that extend the lifespan and oxidative stress resistance of females significantly more than male flies include caloric restriction (Mair et al., 2003; Magwere et al., 2004), Chico mutants (Clancy et al., 2001) and SOD overexpression in motorneurons (Spencer et al., 2003). Until recently, only males have been used in the majority of life extension experiments in *Drosophila*, including reports of *methuselah* (*mth*) mutants (Lin et al., 1998), heat shock protein overexpression (Morrow et al., 2004), and MnSOD mutants (Sun et al., 2002), usually to avoid the confounding factor of female reproduction (Burger and Promislow, 2004). In a study comparing female and male responses to 16 different life-extending interventions, it was found that females fared better than males overall, although in certain situations including hyper-gravity and desiccation resistance, male flies exhibited an increase in lifespan where female flies did not (Burger and Promislow, 2004).

The reasons for these sex biases are unknown, however it is likely to include a combination of diverse factors including reproductive cost (e.g. some interventions may reduce egg laying in females, thereby increasing longevity), sexual selection (male traits making them more attractive to females may reduce their lifespan), sex linkage (a change in gene expression on the Y or X chromosome may be detrimental to males or advantageous to females) or behavioural aspects (Carey et al., 1995; Burger and Promislow, 2004). Reduced reproductive cost can probably be ruled out as a cause of sex bias in this study because the females were fertile: virgin female transgenic stocks crossed to male tubulin-GAL4 flies did produce offspring. It is possible, however, that the female flies had a reduced fecundity. Another possible reason for sex-specific effects on lifespan put forward is a difference in metabolic rates between males and females; female's mitochondria produce less ROS, and have lower levels of oxidative damage than their male counterparts (Vina et al., 2003). The sex-biased phenomenon observed in this study and many others is one that deserves further attention in future studies (Burger and Promislow, 2004).

The positive result of increased oxidative stress resistance in all females overexpressing *DmGS* and/or *Gclc* is tempered by the lack of stress resistance in male flies overexpressing *DmGS* and/or *Gclc*, even though the male flies had higher GSH levels than their female counterparts. Drawing a direct correlation between upregulation of GSH and increased oxidative stress resistance is made difficult by the sex specificity of the effect, despite many other studies reporting a similar sex bias in resistance to oxidative stress.

7 Chapter 7: Oxidative stress resistance and viability in *Drosophila* S2 cells

7.1 Introduction

Glutathione is an important cellular antioxidant that may have an effect on longevity, according to the free radical theory of ageing (Kehrer and Lund, 1994; Meister, 1994; Polekhina et al., 1999; Finkel and Holbrook, 2000; Sohal et al., 2002). Glutathione is maintained by reduction of GSSG by NADH via glutathione reductase, and also by de novo synthesis (Shi et al., 1994). The role of glutathione synthetase (*GS*) was examined in adult *Drosophila melanogaster* in the previous chapter by overexpressing *DmGS* as well as the catalytic subunit of glutamate-cysteine ligase (*Gclc*) and exposing the flies to oxidative stress. These data required an elucidation of the functional role of GS in a simpler system, leading to the use of *Drosophila* cells which under-expressed *DmGS* and which were exposed to treatments causing oxidative stress.

Drosophila Schneider 2 (S2) cells (Schneider, 1972) were used as the cell culture system in this study. S2 cells are derived from late-stage embryos and are thought to have endothelial or haematopoietic origins (Hammond, 2003). The cells are not well differentiated and attach lightly to plastic surfaces, allowing non-enzymatic removal of the cells from flasks (Caplen et al., 2000; Hammond, 2003). The use of the well established S2 cell system has many advantages, including their ease of culture, a doubling time of 20-24 hours (Hammond, 2003), and the data obtained from cell culture studies can potentially be confirmed in whole organism experiments since RNAi has been effective not only in S2 cells but in flies (Giordano et al., 2002; Carthew, 2003; Lee and Carthew, 2003; Morozova et al., 2003; Sledz and Williams, 2005). Gene knockdown by RNAi has been shown to be efficient and reproducible in S2 cells and can be sustained through many cell divisions, provided the protein is not required for viability (Worby et al., 2001).

RNA interference (RNAi) is a naturally occurring gene regulation process responsible for gene silencing using microRNAs (miRNAs) to repress translation of certain transcripts, specifically those involved in development and cell death (Hammond, 2005; Kavi et al., 2005; Sledz and Williams, 2005). The phenomenon of RNAi was first described in *C. elegans* where the presence of exogenous dsRNA resulted in degradation of the complementary mRNA target (Fire et al., 1998). dsRNA has since been shown to induce post-translational gene silencing in a number of species including zebrafish, flies, and mammals (Caplen et al., 2000; Clemens et al., 2000; Birchler et al., 2003). The cellular RNAi mechanism is described in detail in Chapter 1, section 1.3.2.6.

In *Drosophila* cells, the RNAi machinery can be triggered by experimentally introduced dsRNAs. RNAi is used as a reverse genetic technique by degrading endogenous mRNA complementary to the dsRNAs, making it an efficient approach to loss-of-function phenotype analyses (Boutros et al., 2004; Sledz and Williams, 2005). The experimental dsRNAs used in *Drosophila* systems are longer (~500bp) than the ~22bp siRNAs required to achieve silencing in mammalian systems, because in mammals large dsRNA fragments activate an interferon driven inhibition of translation; a process which is absent in *Drosophila* and *C. elegans* (McManus and Sharp, 2002).

In this chapter, the RNAi technique was used to knock down *DmGS* and/or *Gclc* in S2 cells, and in cells exposed to oxidative stress. GCL expression is significantly reduced with age in rats (Liu and Choi, 2000; Zhu et al., 2006), and the concentration of GSH has been shown to decrease with age in many species, including *Drosophila* (Sohal and Weindruch, 1996; Schulz et al., 2000). The capacity of cells to maintain GSH titres during an oxidative challenge has previously been shown to be important in preventing cell death (Shi et al., 1994), and in this study GSH levels and cellular viability in response to *DmGS* and/or *Gclc* knockdown and oxidative challenges were investigated. Furthermore, the effect of the *DmGS* and/or *Gclc*

knockdown on DNA damage following oxidative challenge was determined because GSH depletion has been associated with elevated levels of DNA damage (Reliene and Schiestl, 2005).

7.2 Results

7.2.1 RNA interference in S2 cells

The RNAi technique was used to knock down the transcription levels of *DmGS* and *Gclc* in S2 cells. RNAi in *Drosophila* cells is relatively simple because it is possible to introduce larger (~500bp) portions of RNA molecules directly into cells, which are cleaved into ~22bp siRNA fragments by the cellular machinery. dsRNA templates of 500-700bp are desirable (Kavi et al., 2005), with longer molecules performing more effectively than molecules smaller than 300bp (Caplen et al., 2000; Hammond et al., 2000). Once introduced into the cell, down-regulation of gene expression by siRNA is transient, with the effect noticeable from 18hours post transfection (Caplen et al., 2000), and lasting as long as 6-7 days post transfection (Clemens et al., 2000; Hannon, 2002; Radyuk et al., 2003).

dsRNA was synthesised according to Chapter 2, section 2.23. The protocol for introducing dsRNA into cells used in this study was to soak cells in dsRNA because this has been described as the method of choice for S2 cells (Hammond, 2003). The alternative technique for introducing dsRNA into cells is a calcium phosphate transfection, however this technique is more time consuming, and damaging to the cells. A modified methodology for introducing dsRNA into S2 cells by soaking was devised in this study because soaking the cells in dsRNA alone (Clemens et al., 2000) did not have any effect on *DmGS* and *Gclc* expression levels when tested in this study. The technique used in this study made use of a transfection reagent, and two separate additions of dsRNA, as described in detail in Chapter 2,section 2.24. The use of the cationic lipid transfection agent allowed for the use of lower concentrations of dsRNA

to achieve gene knockdown, and is a non-invasive technique, causing little or no damage to cells (Hammond, 2003). The inhibition of gene expression by dsRNA has been shown to be dose dependent: dsRNA concentrations above 1μg per $5x10^6$ cells plated are reported not to have any additional effect (Caplen et al., 2000). In this thesis, 2μg dsRNA was added to the cells at two time points on the first day, using a total of 4μg (10-17nM) dsRNA for gene knockdown, while in the original paper by Clemens et al (2000), 16μg dsRNA (=40nM for 700bp dsRNA) was added to cells to achieve knockdown (Hammond, 2003).

GFP control

To test whether the process of introducing any dsRNA molecule into S2 cells had an effect on *DmGS* expression, a 500bp dsRNA complementary to a GFP gene segment present on plasmid pQBI25-fC3 (Q-Biogene) was synthesised as described (Chapter 2, section 2.23; AppendixD). The use of GFP as a control for the effect of dsRNA in S2 cells has been established previously (Caplen et al., 2000; Boutros et al., 2004). 4μg (final concentration 10nM) GFP dsRNA was added to S2 cells, and its effect on *DmGS* expression was measured by RT-PCR. Control cells with no dsRNA added had an equivalent volume of transfection agent added with water substituted for the GFP dsRNA.

Figure 7.1 shows that the addition of GFP dsRNA had no significant effect on *DmGS* expression levels over a 2 day time period, which was the proposed duration for the experiments to follow. The time point 'Day 1' was measured 24 hours after the first addition of dsRNA.

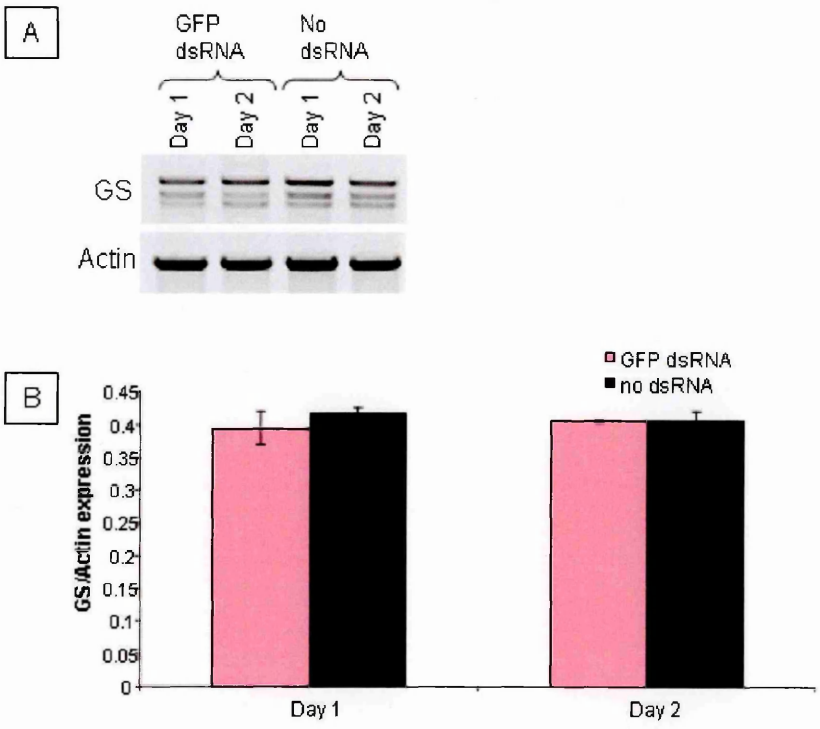


Figure 7.1: RT-PCR data showing *DmGS* expression relative to actin5C. Panel A shows the original agarose gel image of RT-PCR samples of cells with and without the addition of GFP dsRNA. Panel B is a graphical representation of GS:actin expression. There is no statistically significant difference in GS expression between cells treated with GFP dsRNA and no dsRNA, as determined by Students t-test. Means ±SEM are shown.

DmGS and Gclc RNAi

dsRNA corresponding to *DmGS* was synthesised using a primer (Chapter 2, Table 4) from the MRC's Geneservice RNAi construct library (Miah et al., 2004). The primers contain a dual T7 promoter sequence allowing for RNA transcription by T7 transcriptase. The dsRNA sequence corresponded to the final exon of *DmGS*, which all transcripts have in common, ensuring that all known transcripts of *DmGS* would be silenced. *Gclc* dsRNA was synthesised using a primer pair which incorporated a T7 promoter sequence. The dsRNA fragments were 350bp in length for *DmGS*, and 650bp in the case of *Gclc*. The first reaction cDNA products created were sequenced to confirm the identity of the cDNA. 4µg of *DmGS* dsRNA (17nM) or *Gclc* dsRNA (10nM) was then added to S2 cells as described previously, and the expression of each gene was measured for a three day period, starting 24 hours after the first addition of dsRNA. Figure 7.2 shows the knockdown of *DmGS* expression achieved in S2 cells after incubation with *DmGS* dsRNA. The greatest reduction in expression is seen on Day 1 and Day 2, and on

Day 3 a slight increase in the level of *DmGS* transcripts is seen, although the expression is still significantly less than in the control cells which did not have any dsRNA added to them. For the following studies on RNAi, a 2 day assay period was chosen because the transcript knockdown remained low for the first two days following dsRNA incubation.

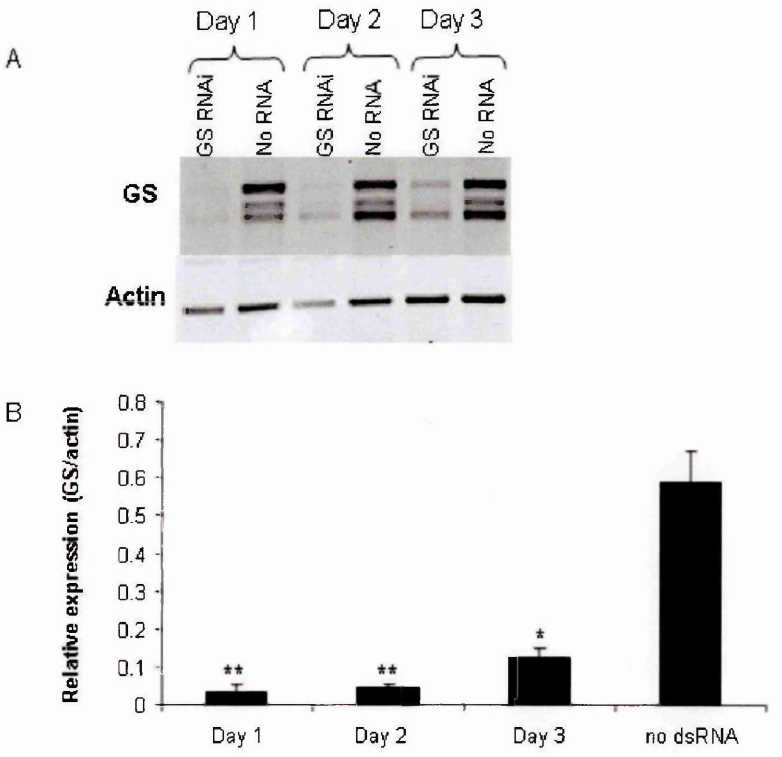


Figure 7.2: *DmGS* mRNA knockdown using dsRNA to *DmGS*. Panel A shows a sample of RT-PCR from cells over a three day period after the addition of *DmGS* dsRNA as well as control cells with no dsRNA added. The results are represented graphically in Panel B where relative expression of *DmGS* is shown. Statistical significance of the difference between *DmGS* expression in control and treatment samples at each time point was calculated using Students 't'test (*P<0.05, **p<0.005). Means ±SEM are shown.

In this study, it was not possible to test GS protein knockdown by western blot because there is no GS antibody available commercially, and the available anti-DmGS polyclonal antibody (Lesley McLellan, pers. comm.) was unreliable when tested on S2 cell extracts.

Gclc expression was also significantly reduced after the corresponding dsRNA was added to the cells (Figure 7.3). Knockdown of the *Gclc* transcript was highly significant over the three day measurement period. The degree of knockdown achieved with the *Gclc* RNAi is more prominent than with *DmGS* RNAi, a gene specific effect which has been noted in several

studies (Caplen et al., 2000). The slight difference in the efficacy of the dsRNA against the two different genes is most likely due to differences in sequence and structure of the siRNAs, and because the *DmGS* dsRNA was almost half the size of the *Gclc* dsRNA (Caplen et al., 2000; Radyuk et al., 2003).

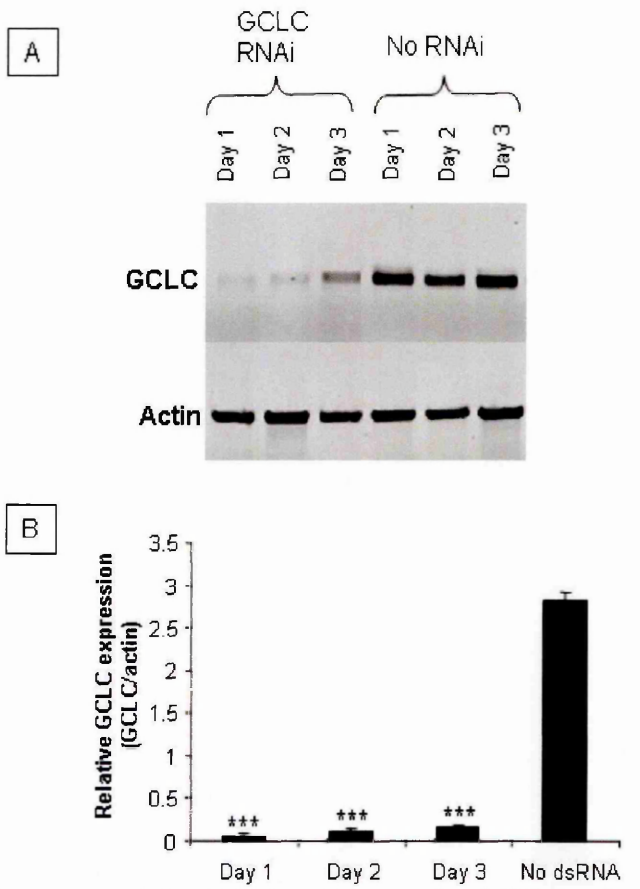


Figure 7.3: *Gclc* mRNA knockdown using dsRNA to *Gclc*. Panel A shows a sample of RT-PCR from cells over a three day period after the addition of *Gclc* dsRNA as well as control cells with no dsRNA added. The results are represented graphically in Panel B where relative expression of *Gclc* is shown. Statistical significance of the difference between *Gclc* expression in control (no dsRNA) and *Gclc* RNAi treated samples at each time point was calculated using Students 't'test (*P<0.001). Means ±SEM are shown.**

7.2.2 Effects of DmGS and Gclc knockdown on cell viability

To determine whether dsRNA transfection *per se* had an effect on cell viability, cells were transfected with GFP dsRNA and viability evaluated by trypan blue exclusion (Chapter 2, section 2.22). The survival of these cells was compared to untransfected cells. Figure 7.4 illustrates that there was no statistically significant difference in viability between cells exposed to GFP dsRNA and control cells over a 2 day period. This experiment suggests that

the presence of dsRNA in S2 cells at the levels used to silence genes (4µg / $3x10^6$ cells) does not have an effect on cell viability.

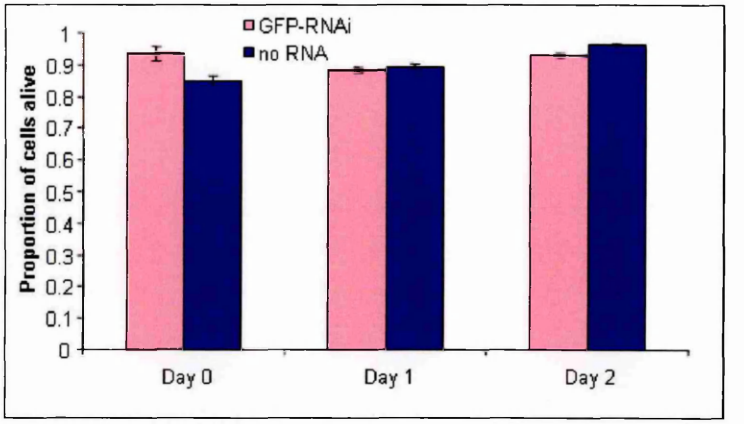


Figure 7.4: Cell viability measured over two days with and without addition of GFP dsRNA. dsRNA was added to the cells at Day 0. Means ± SEM are shown, and means were compared using a Students t-test. No significant difference was found between cells with and without the addition of dsRNA.

Glutathione is an important cellular antioxidant, and it was expected that silencing one or both of the GSH biosynthesis genes would impact on cellular viability. To test whether viability was adversely affected over the two day period following *DmGS* or *Gclc* knockdown, viability was assayed following transfection with dsRNA. There was, however, no difference between the viability of cells exposed to any of the RNAi treatments and the control cells (Figure 7.5).

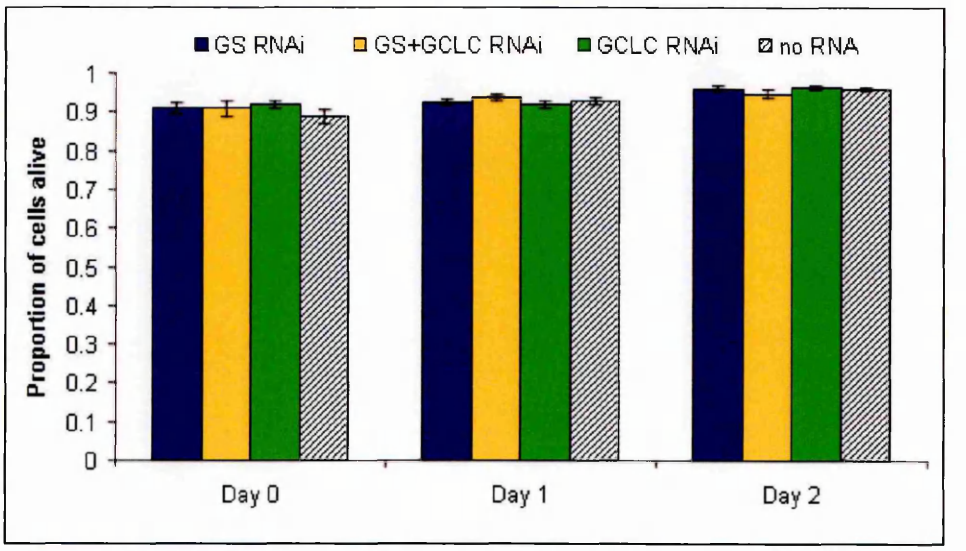


Figure 7.5: Effect of *DmGS* and *Gclc* knockdown alone and the combination of *DmGS* and *Gclc* knockdown on S2 viability. Means ±SEM are shown. Means were compared using the Students t-test, with no significant difference seen between the viability of control cells without dsRNA present, and cells with *DmGS* and/or *Gclc* knockdown.

7.2.3 Comparison of the effects of DmGS and/or Gclc knockdown on glutathione levels in S2 cells

Despite effective RNAi mediated knockdown of *DmGS* and *Gclc* (Figure 7.2 and Figure 7.3), no effect on cell viability was seen in the 48h following transfection. In order to evaluate the impact of *DmGS* and/or *Gclc* knockdown on glutathione levels, transfections with appropriate dsRNAs were carried out and glutathione titres determined 48hours following transfection.

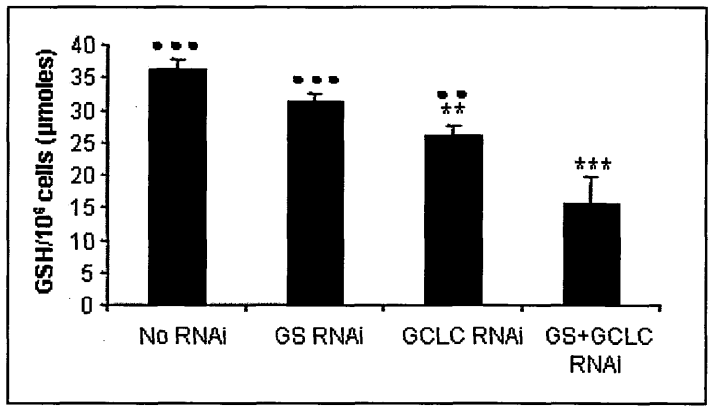


Figure 7.6: Glutathione concentration (µmoles per 10^6 cells) in S2 cells transfected with *DmGS*, *Gclc* or *DmGS+Gclc* RNAi. Samples were taken 48hrs after dsRNA addition. Experiments were performed in triplicate, and repeated twice. Means ±SEM are shown. Statistical significant difference was determined using students 't'test. * indicates a significant difference in GSH levels between RNAi treated cells and those with no RNAi added (p<0.005, ***p<0.001), and • indicates a significant difference between cells treated with the combined *DmGS+Gclc* dsRNA and the remaining RNAi treatments (••p<0.005 •••p<0.001).**

Control cells with no dsRNA added had an average GSH content of 36.3 µmoles per million cells (Table 7.1). Although the addition of *DmGS* dsRNA decreased the GSH content by 13.5% compared to the control, this difference was not statistically significant, as indicated on Figure 7.6. The presence of *Gclc* dsRNA did decrease the GSH content of the cells significantly (by 28.4%) compared to the control cells, while the combination of *DmGS* and *Gclc* dsRNA had the greatest effect on GSH levels. The presence of *Gclc* dsRNA decreased the GSH content of the cells significantly (71.8% of the control), while the combination of *DmGS* and *Gclc* dsRNA had the greatest effect on GSH levels: a reduction to 43% of control levels. These results suggest a cumulative effect of each of the glutathione biosynthesis genes.

Table 7.1: GSH content of S2 cells grown in the presence of *DmGS* and/or *Gclc* dsRNA.

RNAi Treatment	GSH (umoles / 10^6 cells)	% GSH of the control
No RNAi	36.3	-
GS RNAi	31.4	86.5%
GCLC RNAi	26	71.6%
GS+*GCLC* RNAi	15.6	43%

This experiment confirms that RNAi reduces gene expression, rather than completely silencing gene expression (Clemens et al., 2000; Hannon, 2002; Radyuk et al., 2003). There will be a residual level of expression that escapes the RNAi mechanism, especially in genes with a high level of transcription. The residual levels of GSH could also be a result of a slow depletion of the existing cellular GSH pool, implying a long half-life for GS and GCLC proteins, or because *DmGS* and *Gclc* RNAi does not obliterate cellular GSH, but merely reduces the levels considerably.

7.2.4 The effect of chemically induced oxidative stress on S2 cells

Because GSH has an antioxidant role, the effect of oxidative stress in cells with reduced expression of GSH biosynthesis genes would be expected to be greater than in control cells. This hypothesis was tested by inducing conditions of oxidative stress in S2 cells with diethyl maleate (DEM) and hydrogen peroxide (H_2O_2).

DEM is a β-unsaturated carbonyl compound. DEM is conjugated to GSH by an addition reaction, where GSH is added to an activated double bond (Parkinson, 2001). In addition to this chemical reaction, DEM reacts with glutathione enzymatically via glutathione transferases, both of which cause a lowering of endogenous GSH pools (Cardozo-Pelaez et al., 2002). The depletion of cellular glutathione by DEM increases oxidative stress (Back et al., 1998), making DEM a useful chemical test for the ability of S2 cells to withstand oxidative stresses.

H_2O_2 is a normal cellular metabolite which is decomposed by catalase (peroxide produced within peroxisomes) and glutathione peroxidase (peroxide formed in cytosol) (Eklow et al., 1984; Dringen et al., 1999). It is a strong oxidant that is capable of rapidly diffusing through

cell membranes from its endogenous site of origin and is a potent agent of cell death and mutagenic events (Mackay and Bewley, 1989). The glutathione system exerts a key function in the detoxification of H_2O_2 (Figure 7.7), and on exposure to hydrogen peroxide, glutathione oxidation occurs in a matter of minutes (Srivastava et al., 1974; Eklow et al., 1984; Dringen et al., 1999).

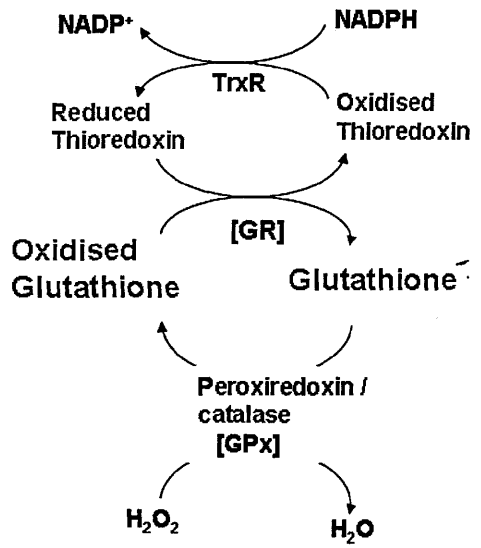


Figure 7.7: Schematic diagram of the role of glutathione in the detoxification of peroxide. Glutathione peroxidase (Gpx) catalyses the reduction of hydrogen peroxide via GSH. In Drosophila, glutathione reductase (GR) does not exist, and instead the thioredoxin system performs the role of glutathione reduction.

The effect of oxidative stress on S2 cell viability

Because both DEM and H_2O_2 expose cells to increased levels of oxidative stress, the hypothesis was that the presence of either of the two chemicals would have a negative effect on viability.

Effect of DEM on cell viability

Two concentrations of DEM were tested; a low dose of 0.25mM and a high dose of 0.5mM. At the low dose of DEM, viability of S2 cells was not affected over the two day period assayed, with no statistical difference between the treated and untreated cells(Figure 7.8). At the higher dose of 0.5mM DEM, however, S2 cells did not have a good survival rate, with 51.4% of the cells alive one day after DEM addition and only 14.1% viability on the second

day following the addition of DEM. There was a significant difference in viability of cells treated with 0.5mM DEM, and those treated with 0.25mM DEM. The lower dose of DEM was chosen for further studies because the dose was not high enough to kill the cells, allowing further analysis to be performed on the cells.

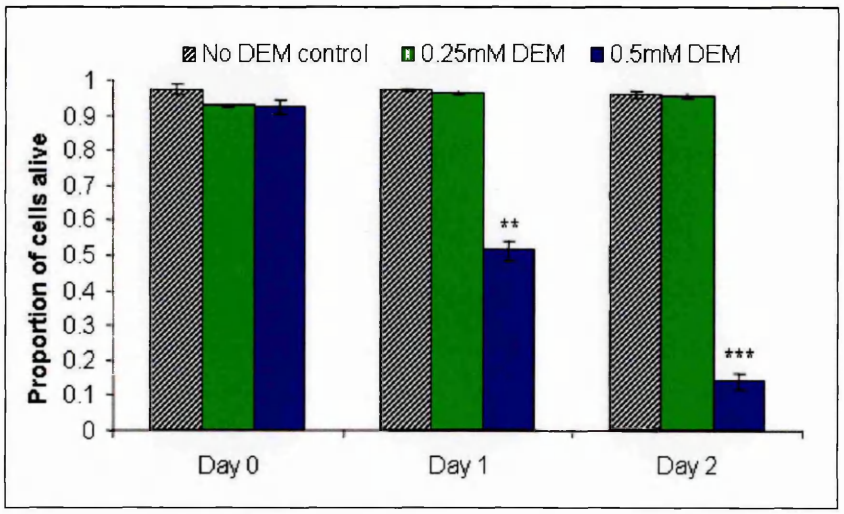


Figure 7.8: S2 cell viability after the addition of 0.25mM DEM (green bars), 0.5mM DEM (blue bars) and a no treatment control (black striped bars). DEM was added to cells on Day 0, and viability was assayed on the two following days. Experiments were repeated at least twice, and performed in triplicate. Mean ±SEM is shown. Means were compared using the students t-test ($p<0.005$, *** $p<0.001$ in comparison with untreated control).**

The effect of H_2O_2 on S2 cell viability

Two concentrations of hydrogen peroxide, 10mM and 30mM, were tested for effects on S2 cell viability. Both concentrations of H_2O_2 had a significant negative effect on viability, with the loss of cell viability more pronounced in cells exposed to the higher concentration of peroxide (Figure 7.9). After 8 hours, the percentage of viable cells was 96.2% in control cells, compared to 89.2% in cells treated with 10mM H_2O_2, and 57.4% in cells treated with 30mM H_2O_2. The viability results obtained are comparable to published data where 70% S2 cell viability was observed after a 4 hour treatment with 20mM H_2O_2 (Radyuk et al., 2001; Radyuk et al., 2006). After 4 hours of treatment, the difference in viability between cells treated with 10mM H_2O_2 and those treated with 30mMH_2O_2 became statistically significant, while the difference in viability between cells treated with 10mM H_2O_2 or 30mM H_2O_2 was statistically significant after 2 hours of treatment, suggesting a rapid effect of peroxide on cell survival.

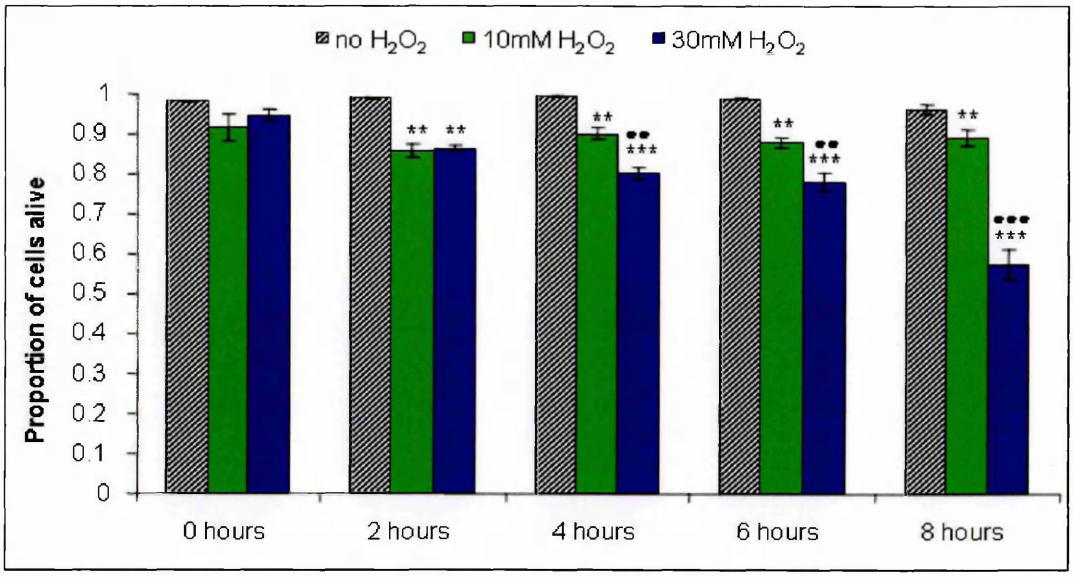


Figure 7.9: S2 cell viability after the addition of 10mM H_2O_2 (green bars), 30mM H_2O_2 (blue bars) and no treatment control (black striped bars). H_2O_2 was added to cells at 0 hours, and viability was assayed at 2 hourly intervals for the period of 1 day. Mean ±SEM is shown. Means were compared using the students t-test, where * indicates a significant difference between the peroxide treated cells and the control cells ($p<0.005$, *** $p<0.001$), and • indicates a significant difference between cells treated with 10mM H_2O_2 and those treated with 30mM H_2O_2.**

The effect of oxidative stress on *DmGS* and *Gclc* expression in S2 cells

The aim of this set of experiments was to test the hypothesis that an increase in environmental oxidative stress would result in an upregulation of glutathione biosynthesis enzymes because GSH is an important cellular antioxidant. *DmGS* and *Gclc* expression levels were measured in S2 cells which had been exposed to 30mM H_2O_2, as well as a low dose (0.25mM) and a high dose (0.5mM) of DEM. *DmGS* and *Gclc* expression was measured 6 hours after H_2O_2 addition, which was the period of the H_2O_2 viability assay in S2 cells described by Radyuk et al (2003). *DmGS* and *Gclc* expression was measured 24 hours after DEM addition, a time point at which cells were viable and not damaged. Figure 7.10 shows the relative expression of *DmGS* (Panels A and C) and *Gclc*(Panels B and D) in S2 cells exposed to one of two different concentrations of DEM, or 30mM H_2O_2.

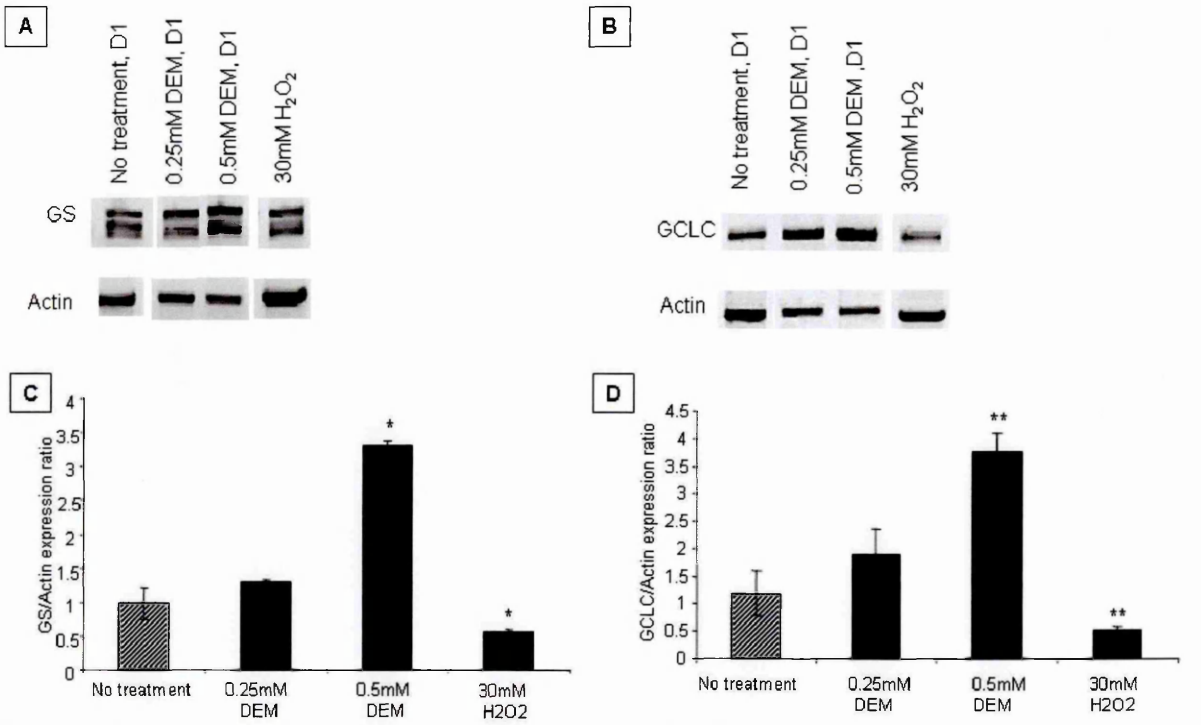


Figure 7.10: *DmGS* expression levels in S2 cells exposed to chemically induced oxidative stress. Panels A and B are examples of agarose gels with *DmGS* (Panel A) or *Gclc* (Panel B) and actin RT-PCR samples. The treatments compared include 0.25mM DEM, 0.5mM DEM, 30mM H_2O_2, and a control of cells grown in SFM media only. Panel C, a graphical representation of the relative intensities of the *DmGS* and Actin bands in Panel A, compares *DmGS* expression in samples treated with 0.25mM, 0.5mM DEM, or 30mM H_2O_2. Panel D is a graphical representation of the gel in Panel B, and compares the relative *Gclc* intensities between cells treated with 0.25mM, 0.5mM DEM, or 30mM H_2O_2. Statistically significant differences between *DmGS* or *Gclc* expression in control cells and treated cells are shown by asterisks (*p<0.05, **p<0.005). Statistical significance was calculated using Students 't'test. Means ±SEM are shown.

The expression of *DmGS* in cells treated with 0.25mM DEM is slightly higher than untreated control cells, although this increase is not statistically significant (Panel C). When cells are treated with 0.5mM DEM, however, the increase in *DmGS* expression is statistically significant. The converse is true for cells treated with H_2O_2, where *DmGS* expression is significantly lower than in untreated cells. The same trends are present when analysing *Gclc* expression in S2 cells treated with DEM and H_2O_2 (Panel D).

This set of experiments shows that enzymes of glutathione biosynthesis are up regulated in the presence of DEM, particularly at the high concentration of 0.5mM. Unexpectedly both *DmGS* and *Gclc* are down regulated by exposure to H_2O_2.

7.2.5 *Effect of DmGS and Gclc RNAi on oxidative stress resistance of S2 cells*

In a previous experiment (Section 7.2.2), it was determined that knockdown of *DmGS* and *Gclc* mRNA, either separately, or in combination, had no effect on S2 cell viability over the time course followed. This led to the current question: whether knockdown of GSH biosynthesis genes would alter viability in S2 cells exposed to oxidative stress.

7.2.5.1 GFP controls

To confirm that the addition of dsRNA *per se* did not have an effect on S2 cell viability in the presence of an oxidative stressor, S2 cells were incubated with GFP dsRNA for 24 hours. The cells were exposed to 0.25mM DEM 24 h after addition of dsRNA, and viability assayed over a two day period.

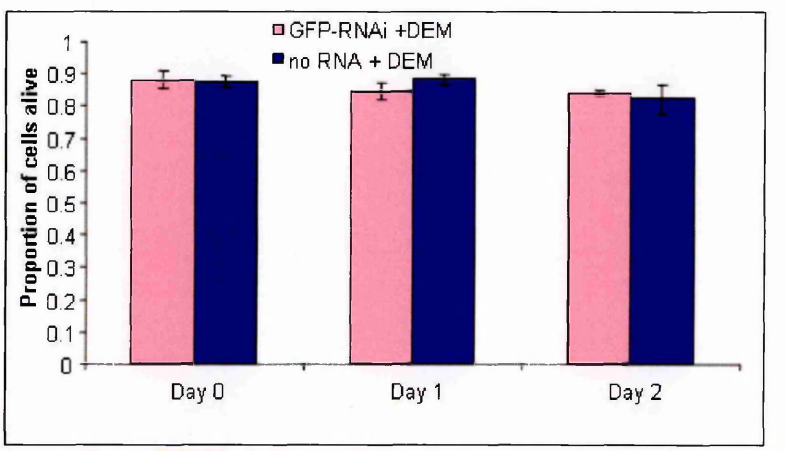


Figure 7.11: Cell viability measured over three days with and without addition of GFP dsRNA, in the presence of 0.25mM DEM. The viability assay was started at the point of DEM addition (one day after the addition of dsRNA). Mean ±SEM is shown, and means were compared using a students t-test. No significant difference was found between cells with and without the addition of GFP dsRNA.

In the presence of 0.25mM DEM, there was no significant difference in the viability of cells treated with GFP dsRNA and those with where there was no dsRNA addition (Figure 7.11). This is corroborated by the experiment on cell viability with GFP dsRNA in the absence of an oxidative stress (section 7.2.1), and a study by Boutrous *et al* (2004), where the presence of GFP dsRNA also had no effect on cell viability.

7.2.5.2 *DmGS* and/or *Gclc* knockdown effect on S2 viability in the presence of DEM

S2 cells were incubated with the relevant dsRNA for 24 hours, followed by an exposure to 0.25mM or 0.5mM DEM for a further 48 hours. Viability was assayed prior to DEM addition and at 24h and 48h following DEM addition. Cells which were not treated with dsRNA were unaffected by 0.25mM DEM treatment over the two day assay time, whereas those cells with *DmGS* knockdown had a statistically significant decline in viability 24 hours and 48 hours after the addition of DEM (Figure 7.12A). The silencing of *DmGS* therefore significantly reduces the DEM-induced oxidative stress resistance of S2 cells. Administration of 0.5mM DEM had significant effects on viability of S2 cells regardless of whether they had been subjected to RNAi mediated knockdown of *DmGS* (Figure 7.12B) . There was a slight decrease in viability between cells with and without *DmGS* knockdown after 24 hours, although this difference was not statistically significant. After 48 hours of DEM exposure, however, there was a statistically significant decrease in viability of cells with *DmGS* RNAi compared to those with no RNAi. The concentration of 0.5mM DEM is likely to be causing a degree of cellular damage too large to be a useful assay in combination with the RNAi treatment. For this reason, the lower dose of 0.25mM DEM was used for further experiments.

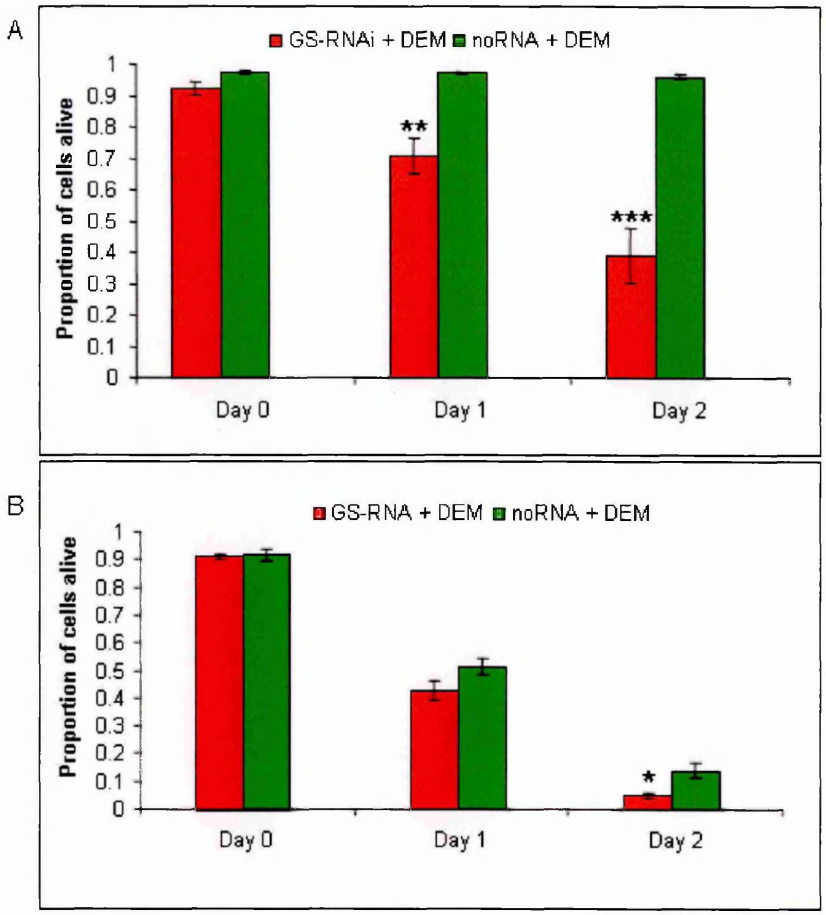


Figure 7.12: Comparison of cell viability after the addition of *DmGS* dsRNA (red bars) and no RNA (green bars). Double stranded RNA was added a day before the addition of DEM on Day 0. Panel A shows viability of cells exposed to 0.25mM DEM, and Panel B shows cell viability after the addition of 0.5mM DEM. Mean ±SEM is shown. Means of cells treated with GS-RNAi were compared to those which were untreated using the students t-test (* $p<0.05$, ** $p<0.005$, *** $p<0.001$).

Figure 7.13 compares the viability of cells treated with a combination of *DmGS* and *Gclc* dsRNA, and cells treated with either *DmGS* dsRNA or *Gclc* dsRNA. The control cells were not exposed to any RNAi treatment, and were unaffected by the addition of 0.25mM DEM over the two day assay period. Cells treated with *DmGS* and/or *Gclc* dsRNA as well as DEM all showed a significant decline in viability over two days. These results suggest that expression of *DmGS* and *Gclc* are equally important in protecting *Drosophila* S2 cells from DEM induced oxidative stress.

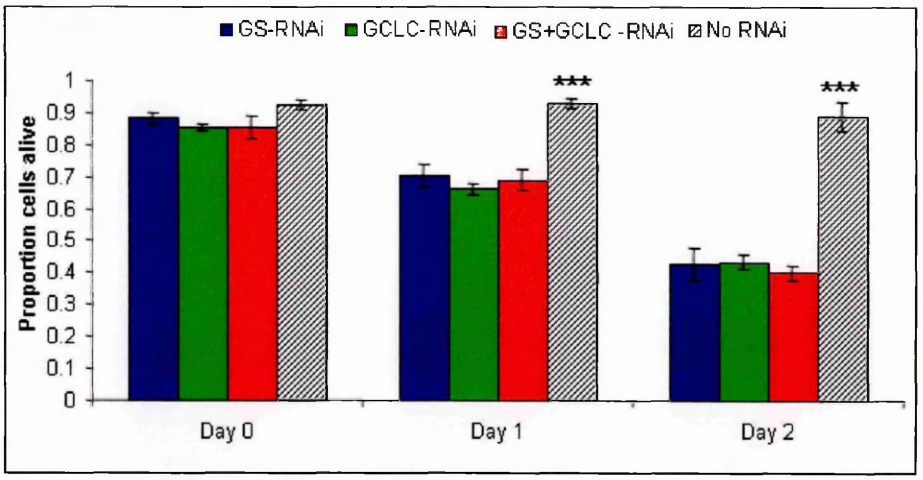


Figure 7.13: Viability of S2 cells with *DmGS* (blue bar) and *Gclc* (green bar) knockdown alone, as well as the combination of *DmGS* and *Gclc* (red bar) knockdown in the presence of 0.25mM DEM. Proportion of cells alive with no dsRNA added are shown with a black striped bar. Mean ±SEM is shown. Means of cells treated with RNAi were compared to those which were untreated using the students t-test (* $p<0.001$ in comparison with the treated samples).**

7.2.5.3 *DmGS* and *Gclc* expression levels in cells treated with *DmGS* or *Gclc* RNAi in the presence of DEM

DmGS and *Gclc* expression were unaffected by 0.25mM DEM, and at 0.5mM DEM, expression of both genes was significantly increased. In order to evaluate whether DEM treatment partially reduced the impact of RNAi-mediated knockdown of *DmGS* or *Gclc*, S2 cells were treated with *DmGS* and/or *Gclc* dsRNA for 24 hours, followed by the addition of 0.25mM or 0.5mM DEM. RNA was extracted from the cells 24 hours and 48 hours after the DEM addition, and RT-PCR was performed on the samples to determine the relative expression of *DmGS* (Figure 7.14) and *Gclc* (Figure 7.15).

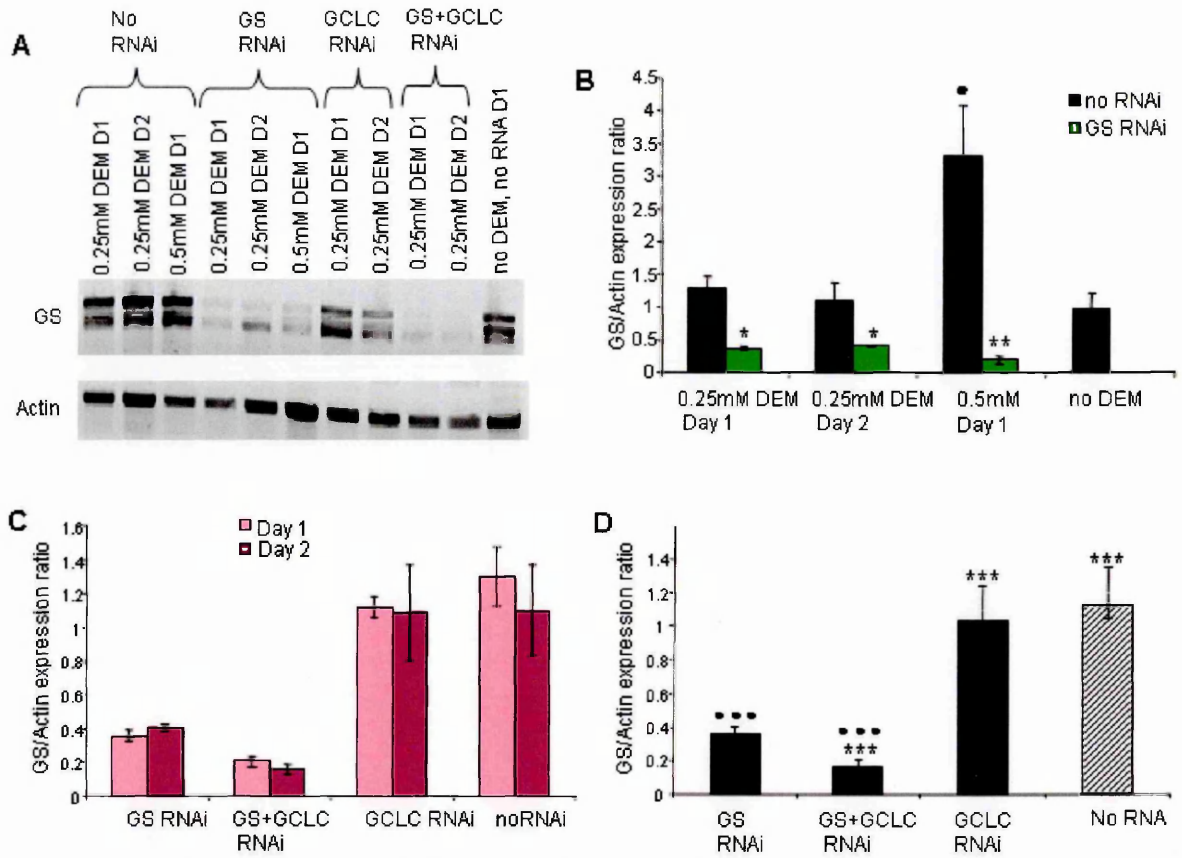


Figure 7.14: *DmGS* expression levels with different RNAi and DEM treatments. dsRNA was added to cells one day before DEM to ensure effective mRNA knockdown. Panel A is an agarose gel with *DmGS* and actin RT-PCR samples. The treatments compared include with and without *DmGS*, *Gclc* or *DmGS*+*Gclc* RNAi under different concentrations of DEM, or no DEM. In the sample labels, D1 indicates samples taken 1 day after DEM addition, and D2 indicates samples taken two days after DEM addition. Panel B compares *DmGS* expression in cells with *DmGS* RNAi (green bars) and those with no RNAi treatment (black bars) at 0.25mM and 0.5mM DEM. Significant differences between the expression of cells with *DmGS* RNAi and no RNAi treatments (*p<0.05, **p<0.005) are shown, as well as the difference in *DmGS* expression between no DEM and different DEM concentrations without RNAi, shown with dots(•P<0.05). Panel C shows *DmGS* expression levels with different *DmGS* and *Gclc* knockdown combinations over two days in the presence of 0.25mM DEM. There was no statistically significant difference of *DmGS* expression levels between Day 1 and Day 2 after DEM addition for any of the RNAi treatments. Panel D shows means of *DmGS* expression levels from Day 1 and Day 2 samples shown in Panel C. Significant differences between the *DmGS* expression levels in the control sample and different RNAi treatments are shown with dots (•••p<0.001), and significant differences between cells treated with *DmGS* RNAi and the other treatments are shown with asterisks (*p<0.001).**

As expected, *DmGS* mRNA was decreased significantly in cells exposed to *DmGS* dsRNA compared to cells without RNAi (Figure 7.14B), thus confirmed the efficacy of the dsRNA in knocking down gene expression. Cells which were exposed to 0.5mM DEM without any RNAi had significantly increased *DmGS* expression compared to the control cells which were not exposed to any DEM or RNAi. In the presence of *DmGS* RNAi, however, this effect was obliterated, showing that the mechanism of the gene silencing could be maintained even as

DmGS was significantly up-regulated. There was no significant difference in *DmGS* expression between cells treated with DEM and *DmGS*/*Gclc* RNAi for 24hours or 48 hours (Figure 7.14C), allowing for the pooling of results from the two days (Figure 7.14D). In the presence of 0.25mM DEM, treatment with *Gclc* RNAi alone had no effect on *DmGS* expression levels, however, cells treated with *Gclc* dsRNA in addition to *DmGS* dsRNA exhibited significantly lower *DmGS* expression than cells treated with *DmGS* dsRNA alone. The reason for *Gclc* dsRNA further decreasing *DmGS* expression levels only when combined with *DmGS* dsRNA is unclear, although it may be possible that *DmGS* expression is upregulated when substrate levels are raised.

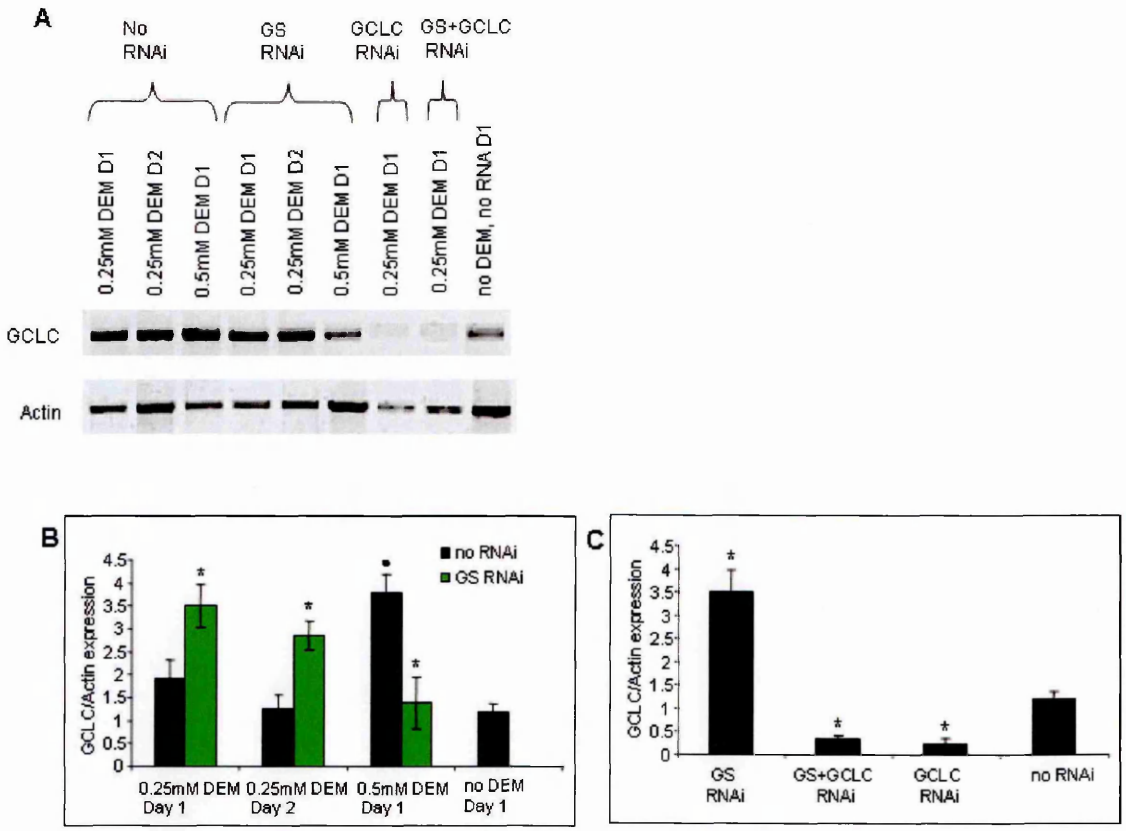


Figure 7.15: *Gclc* expression levels with different RNAi and DEM treatments. dsRNA was added to cells one day before DEM to ensure effective mRNA knockdown. Panel A is an agarose gel with *Gclc* and actin RT-PCR samples. The treatments compared include with and without *DmGS*, *Gclc* or *DmGS*+*Gclc* RNAi under different concentrations of DEM, or no DEM. Panel B compares *Gclc* expression in cells with *DmGS* knockdown (green bars) and those with no RNAi treatment (black bars) at 0.25mM and 0.5mM DEM. Significant differences between the expression of cells with *Gclc* RNAi and no RNAi treatments (*p<0.05) are shown, as well as the difference in *Gclc* expression between no DEM and different DEM concentrations in cells with no RNAi treatment, shown with a dot(•P<0.05). Panel C shows means of *DmGS* expression levels over a two day period. Significant differences between the *DmGS* expression levels in the control sample and different RNAi treatments are shown with asterisks (*p<0.05).

Gclc expression was examined in cells treated with a combination of *DmGS* and/or *Gclc* RNAi, in the presence of DEM. *DmGS* RNAi in cells treated with 0.25mM DEM had the effect of significantly upregulating *Gclc* expression compared to cells with no RNAi treatment (Figure 7.15B). This effect was not a result of the DEM, since untreated cells had *Gclc* expression levels similar to those exposed to 0.25mM DEM. These results suggest that a decrease in *DmGS* expression causes a significant increase in *Gclc* expression at a low dose of DEM. At the higher dose of 0.5mM DEM, however *Gclc* expression was significantly lower in cells treated with *DmGS* RNAi than the control cells with no gene knockdown. This is an interesting result since in the absence of *DmGS* RNAi, *Gclc* expression was found to increase significantly in the presence of 0.5mM DEM, and *Gclc* is also upregulated by *DmGS* RNAi at low DEM concentrations. *DmGS* knockdown combined with a severe oxidative stress has the effect of negating the upregulation of *Gclc* normally seen both in conditions of severe oxidative stress and also with *DmGS* RNAi. The reason for this effect at the high DEM concentration is unknown at this stage, possibly because of general toxicity, or because the high dose of DEM in combination with a removal of *DmGS* transcripts is likely to cause disruptions of GSH signalling pathways.

Cells treated with *DmGS* RNAi only in the presence of 0.25mM DEM showed significantly increased levels of *Gclc* expression compared to control cells with no RNAi treatment (Figure 7.15C). The presence of *DmGS* dsRNA did not have any additional effect on *Gclc* knockdown in cells treated with *Gclc* RNAi. This is noteworthy because the presence of *Gclc* dsRNA did have an additional effect on *DmGS* knockdown in the presence of *DmGS* RNAi (Figure 7.14D).

7.2.5.4 *DmGS* and/or *Gclc* knockdown effect on S2 viability in the presence of H_2O_2

Hydrogen peroxide was used as an alternative source of oxidative stress to compare with the results from the effect of *DmGS* knockdown on oxidative stress resistance using DEM as an

oxidative stressor. S2 cells were treated with *DmGS* dsRNA for 24 hours, followed by an exposure to either 10mM H_2O_2 or 30mM H_2O_2. At both the low dose (Figure 7.16A) and high dose (Figure 7.16B) of H_2O_2 there was no difference in the survival of cells treated with *DmGS* knockdown compared to those without any RNAi treatment. *DmGS* knockdown therefore has no additional effect on cell viability under conditions of peroxide-mediated oxidative stress. It was possible that the lack of an additional effect from *DmGS* RNAi on viability might be specific to *DmGS*, therefore *Gclc* RNAi and a combination of *DmGS* and *Gclc* RNAi was tested in the presence of H_2O_2.

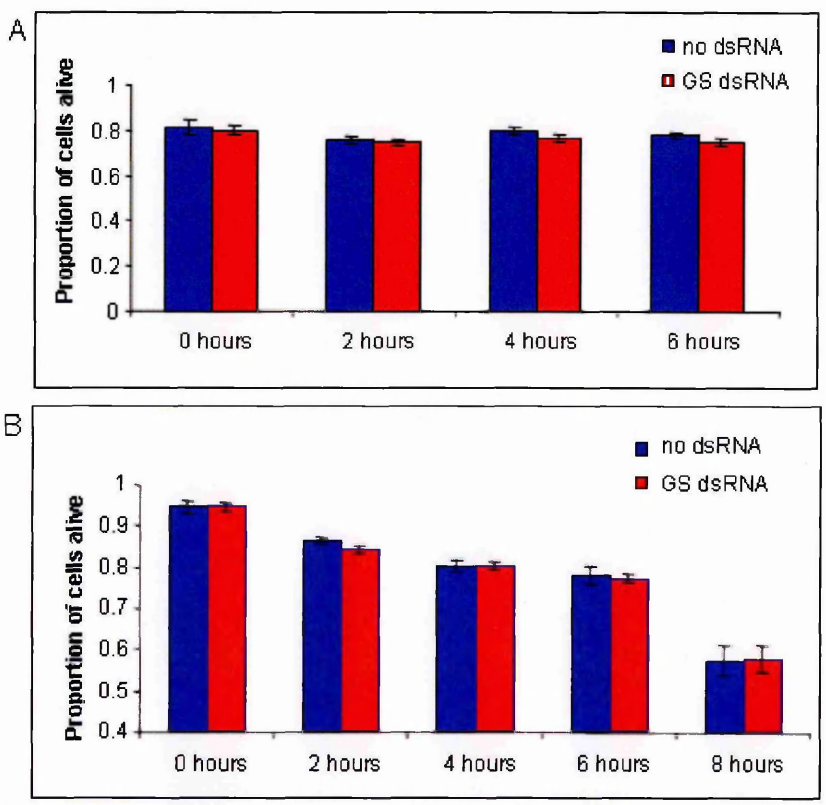


Figure 7.16: Comparison of cell viability after the addition of *DmGS* dsRNA (red bars) and no RNA (blue bars). Double stranded RNA was added a day before the addition of H_2O_2 on Day 0. Panel A shows viability of cells exposed to 10mM H_2O_2 and Panel B shows cell viability after the addition of 30mM H_2O_2. Mean ±SEM is shown. Means were compared using the Student's t-test, and no significant difference was found between survival of cells treated with GS dsRNA and those not treated with any dsRNA .

The following experiment aimed to determine whether the *DmGS* and *Gclc* double knockdown would have an effect on the viability of cells exposed to 10mM H_2O_2. As illustrated in Figure 7.17, there was no statistically significant difference between the viability of cells treated with

DmGS dsRNA, *Gclc* dsRNA or a combination of the two. None of the treatments had any effect on viability in the presence of peroxide, as there was no decrease in viability compared to control cells which did not have any RNAi treatment.

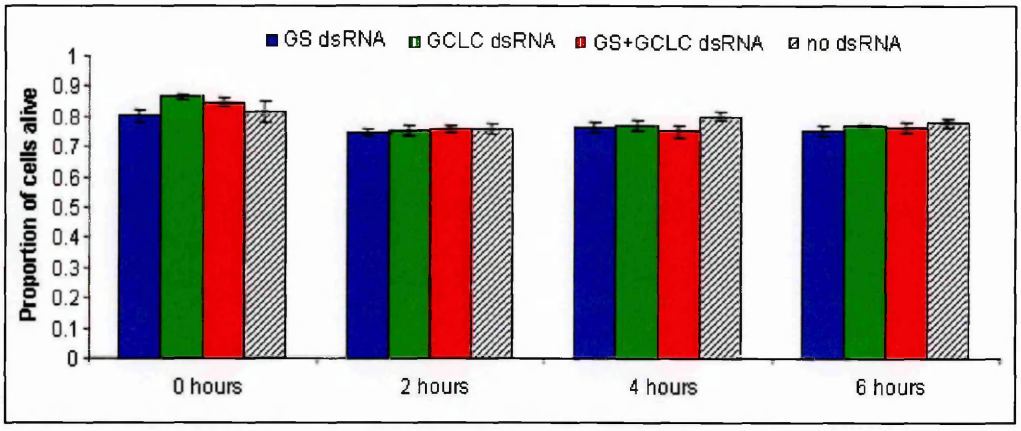


Figure 7.17: Viability of S2 cells with *DmGS* (blue bar) and *Gclc* (green bar) knockdown alone, as well as the combination of *DmGS* and *Gclc* (red bar) knockdown in the presence of 10mM H_2O_2. Proportion of cells alive with no dsRNA added are shown with a black striped bar. Mean ±SEM is shown. Means were compared using the Students t-test (no significant differences were found between the different treatment groups over 6 hours).

These results are interesting in light of the negative effect of *DmGS* and/or *Gclc* knockdown on cells exposed to a low concentration of DEM (Section 7.2.5.2). The difference between the H_2O_2 and DEM treatments on cells treated with *DmGS* and/or *Gclc* dsRNA is most likely due to the glutathione depleting effects of DEM, which is examined further in the discussion section of this chapter. H_2O_2 is primarily detoxified by catalase, which may also explain the apparent lack of effect of GSH depletion on peroxide resistance.

7.2.6 Glutathione levels in cells treated with DmGS and/or Gclc dsRNA, and DEM

Glutathione levels in cells treated with *DmGS* and/or *Gclc* dsRNA in the presence of 0.25mM DEM were calculated to determine the cumulative effect of DEM and *DmGS*/*Gclc* RNAi on reducing cellular glutathione. The appropriate dsRNA was added to samples and incubated for 24 hours, after which time 0.25mM DEM was added to the cells. After a further 24 hours, samples were taken for the glutathione assay.

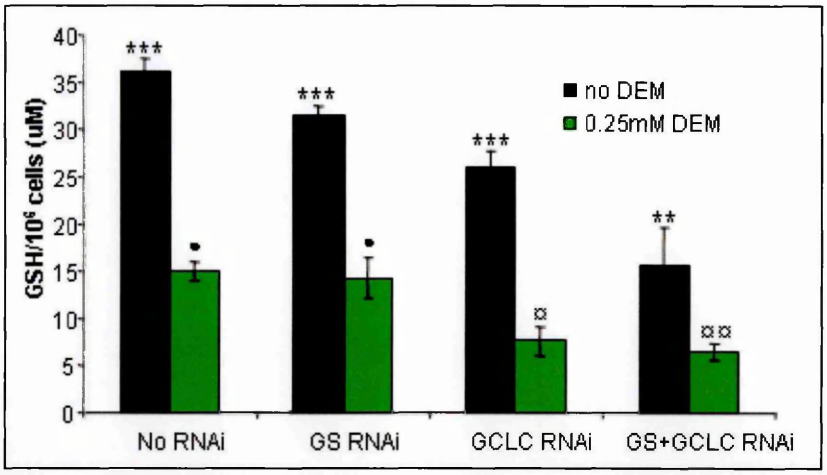


Figure 7.18: Glutathione concentration (µM per 10^6 cells) in S2 cells with *DmGS*, *Gclc* or *DmGS+Gclc* RNAi. dsRNA was added to cells at 0hrs, 0.25mM DEM was added at 24hrs, and samples for GSH determination were taken at 48hrs. Experiments were done in triplicate, and repeated twice. The graph shows a comparison of GSH in cells with (green bars) and without (black bars) the addition of DEM. Means ±SEM are shown. Statistically significant differences were determined using student's t'test. * indicates a statistically significant difference between GSH levels in cells grown in the presence of DEM and cells grown without DEM in presence of various dsRNAs (p<0.005, ***p<0.001). • indicates a significant difference in GSH levels between cells treated with *GS+Gclc* RNAi, and all the other RNAi treatments, in the presence of DEM (•p<0.05). ¤ indicates a significant difference between cells not treated with any dsRNA, and those treated with *GS* and/or *Gclc* dsRNA in the presence of DEM.**

The presence of DEM was shown to significantly lower the level of GSH in the presence of *DmGS* or *Gclc* RNAi, as well as in control cells with no RNAi (Figure 7.18). DEM lowered cellular GSH concentrations to a greater extent than *DmGS* RNAi or *Gclc* RNAi did, except in the case of a combination of *DmGS* and *Gclc* RNAi, which reduced GSH to the same GSH concentration (~15 µmoles/10^6 cells) as the DEM treatment did (Table 7.2). The lowest GSH content (6.5 µmoles/10^6 cells) was in cells treated with DEM and *Gclc* RNAi (Table 7.2), which caused a decrease in GSH to 18% of untreated control cells. The presence of *DmGS* RNAi in cells exposed to DEM had no additional effect on GSH levels. This is consistent with GSH results obtained for cells treated with *DmGS* RNAi in the absence DEM, described in section 7.2.3

In the presence of DEM, there was no statistical significant difference in GSH levels between cells treated with *Gclc* RNAi and cells treated with a combination of *DmGS* and *Gclc* RNAi (green bars in Figure 7.18). This is in contrast to cells that had not been treated with DEM,

where *DmGS* and *Gclc* knockdown had an additive effect, decreasing the cellular GSH levels (Figure 7.6). The reason for this difference could be that the combination of DEM and *Gclc* RNAi reduces the GSH content (6.5 μmoles/10^6 cells) to a level beyond which cell survival is not supported, ie, GSH levels could not be further reduced in viable cells. The combination of DEM treatment and *DmGS* and/or *Gclc* RNAi resulted in ~50% reduction in cell viability compared to untreated cells (Figure 7.13). This observation is consistent with evidence in *S. cerevisiae* that approximately 15% of the wild type glutathione content is critical for survival (Kistler et al., 1986; Ullmann et al., 1996).

Table 7.2: GSH content of S2 cells grown in the presence of *DmGS* and/or *Gclc* dsRNA and 0.25mM DEM. Data in this table originates from the experiment represented in Figure 7.18.

	NO DEM		0.25mM DEM	
RNAi treatment	**GSH (μmoles/10^6 cells)**	**%GSH of untreated control**	**GSH (μmoles/10^6 cells)**	**%GSH of untreated control**
No RNAi	36.3	-	15	41.3%
GS RNAi	31.4	13.5%	14.3	39.4%
Gclc RNAi	26	28.4%	7.6	20.9%
GS+*Gclc* RNAi	15.6	57%	6.5	17.9%

7.2.7 Effect of DmGS and Gclc knockdown on DNA damage

Assays of DNA fragmentation have been used to evaluate DNA damage in cells exposed to oxidative stress conditions (Radyuk et al., 2003; Nygren et al., 2005; Reliene and Schiestl, 2005). In the following experiments, S2 cells exposed to DEM-mediated oxidative stress and *DmGS/Gclc* knockdown were assayed for DNA damage using the Comet assay. In the Comet assay, also known as the single-cell gel (SCG) cells are immobilised and subjected to electrophoresis, with the effect that the DNA of cells with an increased frequency of DNA double strand breaks migrates further towards the anode (Tice et al., 2000). The current Comet assay method of choice in genotoxicology studies employs alkaline (pH >13) electrophoresis (Singh et al., 1988), which causes DNA denaturation and unwinding due to hydrogen bond

disruption, and the weakened bonds are transformed to single strand breaks (Olive and Banath, 1993; Tice et al., 2000).

Because the Comet assay requires a minimum of 75% cells alive to prevent false positives, the cells analysed were exposed to 0.25mM DEM only, and not the higher concentration of 0.5mM. Cells were analysed after a 24 hour incubation with the relevant dsRNA, followed by a further 24 hour exposure to 0.25mM DEM. The Comet protocol used in this study was a combination of various methodologies, and is described in detail in Chapter 2, section 2.26. Briefly, cells were embedded in agarose on a glass slide, then lysed and treated with an alkaline solution to denature the DNA and hydrolyse sites of damage. The embedded cells were then placed in an electrophoresis tank at 20 volts and stained with SYBR Green to reveal the DNA. The samples were then visualised by epifluorescence microscopy, and analysed using the CometScore software application described below.

7.2.8 Comet analysis using CometScore v1.05 software application

The CometScore freeware application, available from http://autocomet.com, uses image processing algorithms to analyse specific features of the tail and head of the DNA 'comet'. The image of the cell obtained with fluorescent microscopy is analysed by first drawing a box around the whole head and tail, thus separating the comet from the background, followed by designating the central portion of the head (Figure 7.19). The intensity of each pixel in the image is directly proportional to the amount of DNA at that location in the cell, allowing for a calculation of the amount of DNA in the designated area.

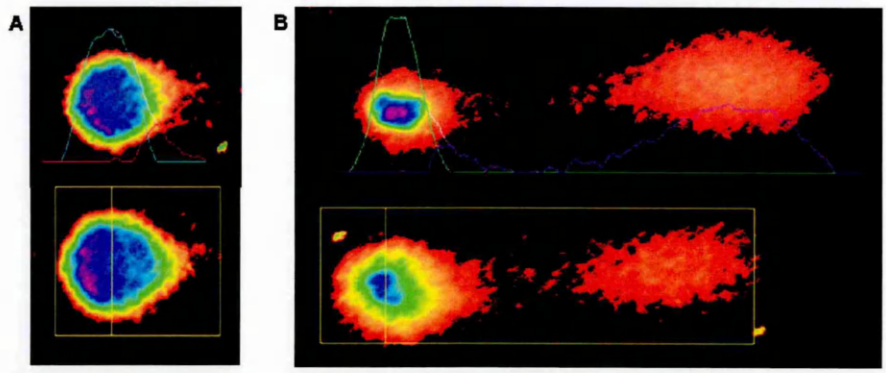


Figure 7.19: Images from the CometScore program. First a box is drawn around the total comet area, followed by clicking on the centre of the nucleus. The program then gives a graphical representation of the DNA amount in the nucleus (green line), and DNA in the tail (purple line). Panel A shows two cells with little DNA damage, and small tail lengths. Panel B shows two cells with a greater extent of DNA damage exhibiting the longer tail length.

The parameter used in this study to evaluate the extent of DNA damage in comet analysis is the percentage DNA in the tail, because this is linearly related to DNA break frequency up to about 80 % DNA in the tail. The intensity of the comet tail relative to the head reflects the number of DNA breaks, because the DNA loops containing a break lose their supercoiling and become free to extend toward the anode, or small broken fragments can leave the nucleus and travel towards the anode (Collins, 2004).

At least 150 cells from each treatment in this study were analysed in duplicate, giving data on numerous cells which tended to have a large range of values for % DNA in the tail. Each experiment was performed at least twice. Tice et al. (2000) suggest classifying the comet data into five categories based on the % DNA in the tails, allowing the visualisation of the average extent of DNA migration in a culture of cells. In this study, data were broken into five classes, based on the %DNA in the tail of each cell: 0-19%, 20-39%, 40-59%, 60-79%, and 80-100%, as illustrated in Figure 7.20 and Figure 7.21.

7.2.9 GFP RNAi control

Cells were exposed to GFP dsRNA to determine whether the presence of dsRNA molecules would interfere with the DNA repair process, making cells more susceptible to DNA damage. The GFP dsRNA was added in the absence (Figure 7.20A) and presence (Figure 7.20B) of

DEM. In both cases, the addition of the GFP dsRNA did not cause any significant change in the amount of DNA damage. The presence of DEM did, however, cause an increase in the proportion of cells with a higher percentage DNA in the tails, indicating a higher level of DNA damage. The significant effect of DEM on DNA fragmentation was observed both in the presence and absence of GFP dsRNA, and is seen clearly when the mean percentage DNA in the tails are compared in Figure 7.20C, where DEM increases the DNA in the tail from 18% to 28%.

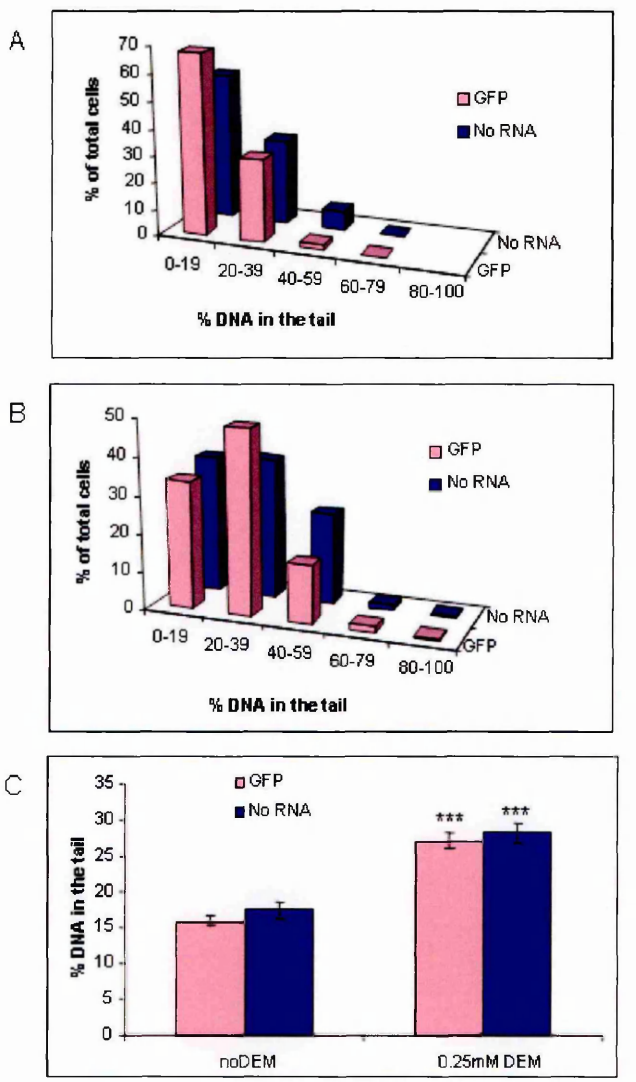


Figure 7.20: The effect of GFP dsRNA and DEM on DNA fragmentation (Percentage DNA in the tail). The analysis was performed on data obtained from the Comet assay. Cells were treated with no DEM (Panel A), or with 0.25mM DEM (Panel B). Panel C shows a comparison of the means from each group. Means ± SEM are shown. Means were compared using Student's t-test. There was no significant difference between the noRNA and the GFP control groups, however there was a highly statistically significant difference between the means of cells treated with DEM, and those not treated with DEM (*p<0.001).**

7.2.10 The effect of DmGS and/or Gclc RNAi on S2 cell DNA damage in cells exposed to an oxidative stress

This set of experiments aimed to determine the effect of *DmGS* and/or *Gclc* RNAi on DNA damage, and whether *DmGS* and/or *Gclc* knockdown had an additional effect to DEM. Cells were treated with dsRNA and DEM as described in the previous section.

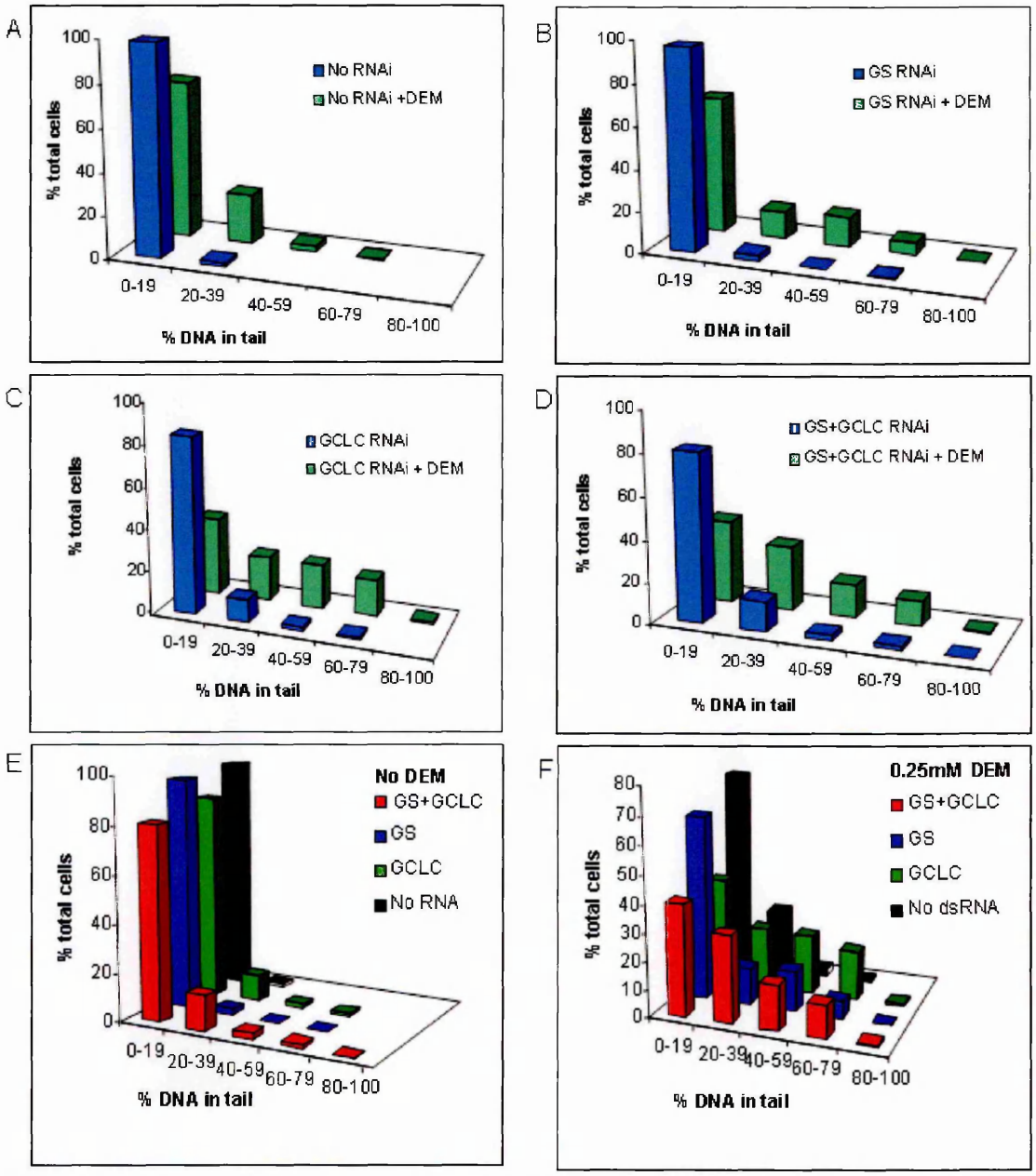


Figure 7.21: The effect of *DmGS* and/or dsRNA and DEM on DNA fragmentation. The analysis was performed on data obtained from the Comet assay. Cells were treated with no DEM (blue bars in panels A and C) or with 0.25mM DEM (green bars in Panels B and D). Cells were treated with either no RNAi (Panel A), *DmGS* RNAi (Panel B), *Gclc* RNAi (Panel C) or a combination of *DmGS* and *Gclc* RNAi (Panel D). Panel E compares the DNA damage in cells treated with combinations of RNAi in the absence of DEM, and Panel F compares DNA damage of cells treated with 0.25mM DEM, as well as *DmGS* and *Gclc* RNAi.

The percentage DNA in the tails of cells treated with no RNAi (Figure 7.21A), *DmGS* RNAi (Figure 7.21B), *Gclc* RNAi (Figure 7.21C) and a combination of *DmGS* and *Gclc* RNAi (Figure 7.21D) increased in the presence of 0.25mM DEM. When comparing the effect of the RNAi combinations in the presence of DEM (Figure 7.21F) and in the absence of DEM (Figure 7.21E), it appears that the combination of *DmGS* and *Gclc* RNAi caused the most DNA damage, with the effect more apparent in cells subjected to the DEM treatment. Figure 7.22 clearly illustrates the effect of DEM on DNA damage by comparing the means of the percentage DNA in the tail of the various treatments. Cells exposed to DEM always displayed a significantly greater degree of DNA damage compared to their untreated controls, regardless of the accompanying RNAi treatment. In the absence of DEM, there was no significant difference in percentage DNA in the tail between cells treated with no RNAi, and those treated with *DmGS* and/or *Gclc* RNAi (black bars in Figure 7.22). This suggests that the knockdown of *DmGS* or *Gclc* or the combination of the two does not increase the incidence of DNA damage in S2 cells. In the presence of 0.25mM DEM, however, there was a statistically significant increase in the amount of DNA in the tail of cells treated with *Gclc* RNAi, with or without *DmGS* RNAi, compared to cells not treated with any RNAi (Figure 7.22). In the presence of DEM, the addition of *DmGS* and *Gclc* dsRNA did not have a cumulative effect on DNA damage compared to *Gclc* dsRNA alone. Cells exposed to *DmGS* RNAi in the presence of DEM had a 6% increase in DNA in the tail compared to DEM-treated cells with no RNAi, although this increase is not statistically significant. *DmGS* RNAi, therefore, does not have any effect on DNA damage in the presence or absence of DEM, while *Gclc* RNAi does have an effect on DNA damage in the presence of the oxidative stressor DEM.

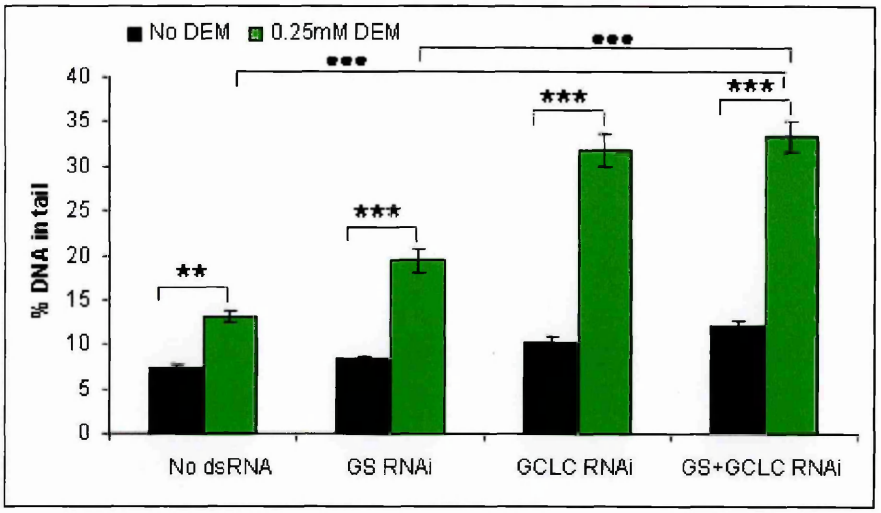


Figure 7.22: Percentage DNA in tail in the presence of DEM or not, and with different combinations of *DmGS* and *Gclc* knockdowns. The means compared originate from the data presented in Figure 7.21. Asterisks indicate a statistically significant difference (p<0.01, ***p<0.001) between DEM and no DEM treatment. ●●● indicates a statistical difference (p<0.001) between *DmGS*+*Gclc* treated cells and other treatment groups with DEM. Statistical significance was calculated using Student's t-test. Means ±SEM are shown.**

7.3 Discussion

The use of RNAi in *Drosophila* cell culture has previously been shown to be a quick and effective method to characterize the role of specific genes by causing significant down-regulation of target genes (Radyuk et al., 2003). In work reported in this chapter, RNAi was used to decrease the expression of *DmGS* and *Gclc* in S2 cells in an attempt to elucidate the role of *DmGS* in *Drosophila* cell viability, resistance to oxidative stress and protection from oxidative stress induced DNA damage. The use of RNAi knockdown of *DmGS* and *Gclc* allows the study of GSH depletion without the possibility of unrecognized effects of GSH depleting chemicals.

In order to fully characterize *DmGS* in S2 cells, it was also necessary to study the effects of *Gclc* knockdown, and a combination of *DmGS* and *Gclc* RNAi since *DmGS* catalyses only one of the two steps of GSH biosynthesis. RNAi has been used in the S2 cell culture system in previous studies, eliciting a robust gene silencing effect which lasted up to 7 days (Clemens et al., 2000; Hannon, 2002; Radyuk et al., 2003). The results presented in this chapter show that

RNAi resulted in a significant reduction of *DmGS* and *Gclc* expression, both individually and in combination. 1.3μg dsRNA was added per million cells (~10-17nM, depending on the length of the dsRNA fragment), resulting in a significant reduction of the relevant transcript after 24 hours, with the effect lasting for at least 3 days. The two day time period was thought to be sufficient to measure GSH biosynthesis gene knockdown since the half life of GSH is 5 hours in rat cell culture (Komlosh et al., 2002), and lasts only 2 minutes in human serum (Lomaestro and Malone, 1995).

DmGS knockdown was slightly less efficient than *Gclc* knockdown, most likely due to the smaller size of the *DmGS* dsRNA (350bp) compared to the *Gclc* dsRNA (650bp). It has been previously reported that the optimal length of dsRNA for efficient RNAi in S2 cells is in the region of 500-700bp (Hammond et al., 2000; Kavi et al., 2005), with 300bp being the lower limit for efficient knockdown (Caplen et al., 2000). The difference in the secondary structure of dsRNA for different genes has also been shown to be responsible for the disparity in knockdown efficacy of certain genes (Caplen et al., 2000; Radyuk et al., 2003).

The effect of oxidative stress on control S2 cells was evaluated. The hypothesis tested was that cells exposed to oxidative stress would have decreased GSH levels, decreased viability, increased DNA damage and increased *DmGS* and/or *Gclc* gene expression. To test the response of cells to oxidative stress, different concentrations of DEM and H_2O_2 were used to induce conditions of oxidative stress. DEM-induced oxidative stress has previously been shown to cause a decrease in human cell viability (Back et al., 1998), most likely through induction of apoptosis (Valverde et al., 2006). It was found that both DEM and H_2O_2 caused a decrease in S2 cell viability. Although a low dose of DEM did not have a statistically significant effect on cell viability, cells experienced a 58% decrease in GSH levels, and a 10% increase in DNA damage. This phenomenon of a significant reduction in cellular GSH having no immediate effect on viability has been noted in mouse cells, where cells survive with as

little as 2% of the wild type GSH levels, although at those levels cells do not grow as well as wild type cells (Valverde et al., 2006).

It is interesting to note that the response of *DmGS* and *Gclc* expression to the two oxidants was not the same, as summarized in Table 7.3. Exposure to DEM caused an upregulation of *DmGS* as well as *Gclc*, while cells exposed to H_2O_2 had reduced expression of the two genes. The increase in *DmGS* and *Gclc* expression was only statistically significant at 0.5mM DEM, although the trend was present at the lower dose of 0.25mM DEM. In previous studies using rat and human tissue culture systems, exposure to H_2O_2 was shown to upregulate *Gclc* and increase cellular GSH content (Soltaninassab et al., 2000; Seo et al., 2004). In *S. cerevisiae, Gclc* was upregulated in response to H_2O_2 (Monje-Casas et al., 2004), although there are conflicting reports of the *GS* expression upon exposure to H_2O_2; one study saw a slight reduction in *GS* expression (Monje-Casas et al., 2004), while another showed an increase in *GS* expression (Carmel-Harel and Storz, 2000). The conflicting *GS* and *Gclc* expression levels in response to DEM compared to H_2O_2 is likely a result of the GSH depleting effects of DEM which would cause an increase in GSH biosynthesis gene expression, while exposure to H_2O_2 would have an indirect effect on GSH by increasing the oxidative stress in the cell, however catalase is the major H_2O_2 detoxifier.

DEM was chosen as the source of oxidative stress for further experiments because the results could be compared to DEM-induced oxidative stress in adult flies, and DEM also has the added advantage of being a glutathione depletor, allowing the examination of conditions of severe glutathione depletion in S2 cells when combined with *DmGS/Gclc* RNAi. DEM-induced GSH depletion is well established as a source of oxidative stress, which has been shown to cause an increase in free radical generation after 12 hours (Back et al., 1998). A withdrawal of GSH leads to a fall in mitochondrial dehydrogenase, an increase in ROS

production, a rise in pro-apoptotic RNA, and a fall in anti-apoptotic RNA, and finally after approximately 48 hours, DNA fragmentation occurs (Shen et al., 2005; Valverde et al., 2006).

Table 7.3: Summary of principal results obtained with various RNAi and oxidative stressors. Up arrows indicate and increase, down arrows indicate a decrease, and horizontal arrows indicate no change. Statistically non significant changes are indicated by (N-S) .

RNAi treatment	Oxidative stressor	Effect
None	30mM H_2O_2	↓ viability by 50% ↓ *GS* expression ↓ *GCLC* expression
	0.25mM DEM	↔ viability ↑ DNA damage by 10% ↓ GSH by 58% ↑ *GS* expression (N-S) ↑ *GCLC* expression (N-S)
	0.5mM DEM	↓ viability by 50% (after 1 day) ↑ *GS* expression ↑ *GCLC* expression
GS	None	↔ viability ↔ DNA damage ↓ GSH by 13.5% (N-S)
	30mM H_2O_2	↔ viability
	0.25mM DEM	↓ viability by 50% (compared to no RNAi) ↑ DNA damage (N-S) ↔ GSH (compared to no RNAi) ↑ *GCLC* expression
GCLC	None	↔ viability ↔ DNA damage ↓ GSH by 28.4%
	30mM H_2O_2	↔ viability
	0.25mM DEM	↓ viability by 50% (compared to no RNAi) ↑ DNA damage (compared to no RNAi) ↓ GSH by 79.1% ↔ *GS* expression (compared to no RNAi)
GS + *GCLC*	None	↔ viability ↔ DNA damage ↓ GSH by 57%
	30mM H_2O_2	↔ viability (compared to no RNAi)
	0.25mM DEM	↓ viability by 50% (compared to no RNAi) ↑ DNA damage (not more than *GCLC* RNAi) ↓ GSH by 82.1% (N-S more than *GCLC* RNAi) ↓*GS* expression compared to GS RNAi ↔ *GCLC* expression compared to *GCLC* RNAi

The next hypothesis tested was that a decrease in either or both of the glutathione biosynthesis genes would decrease glutathione levels, increase DNA damage, and have an adverse effect on

cell viability. It was found that RNAi mediated knockdown of *DmGS* had little measurable effect on viability, DNA damage or GSH levels. Similarly, *Drosophila* SOD (Seto et al., 1990) and catalase (Mackay and Bewley, 1989) mutants expressing 50% of wild type enzyme levels have viability comparable to wild type strains. In addition, GS null human fibroblasts subjected to a Comet Assay were not found to be more sensitive to DNA damage than control cells, despite lower GSH levels, leaving the authors to conclude that GSH is not essential to protect DNA against DNA damage. It is, however, unlikely that the *DmGS* mRNA remaining after the RNAi treatment were active throughout the period of the assay, as the half-life of *GS* mRNA, measured in *S. cerevisiae,* is 16 minutes (Monje-Casas et al., 2004), and the protein half life, measured in *E. coli,* is 50 minutes (Inoue et al., 1996). Another possibility for the lack of effect is that GS is dispensable in *Drosophila*, since GS has been found to be dispensable for viability in *S. cerevisiae*, and *E. coli* (Grant et al., 1997; Cardozo-Pelaez et al., 2002). In humans, however, homozygous *GS* null mutations tend to be lethal, although heterozygous *GS* mutation carriers with *GS* expression of only 50% of control levels display no symptoms (Janaky et al., 1999). It is likely that the reduction of *DmGS* by RNAi left enough *DmGS* to be translated to carry out baseline GSH synthesis.

While *DmGS* RNAi did decrease the cellular GSH concentration, it was a statistically non-significant change, however *DmGS* RNAi in combination with *Gclc* RNAi reduced the GSH content of cells significantly more than *Gclc* alone, indicating that *DmGS* does indeed play an important role in GSH synthesis. Overexpressing *DmGS* and *Gclc* in rats increased GSH synthesis to a greater degree than *Gclc* alone, while upregulating *GS* in a human liver cell line also resulted in a significant increase in GSH, despite no change in GCL expression (Huang et al., 2000). Reduction of GS activity has also been shown to have lowered GSH content when there is no change in GCL activity, as observed post surgery in humans (Luo et al., 1998), and in *Tat* mice (Choi et al., 2000). It may be the case that GS is able to increase Gclc activity,

since it appears that GS has an effect on GSH content only in the presence of Gclc. In addition, when GSH levels are decreased, the feedback inhibition of GCL is released, thus requiring an increase in GS to produce more GSH. This release of inhibition may also explain why a drop in *DmGS* expression has a greater effect on GSH concentration in the absence of *Gclc*.

Gclc knockdown decreased cellular GSH concentration by a significant 28.4%, although there was no change in viability of the cells, or in the extent of DNA damage, compared to untreated control cells. A combination of *DmGS* and *Gclc* knockdown also had no effect on viability or DNA damage within the cells, although cellular GSH was reduced dramatically by 57%. Knockdown of *DmGS* and/or *Gclc* transcription, which may cause a severe decline in GSH concentration, therefore does not have an effect on cell viability in standard culture conditions. The final set of experiments aimed to investigate whether the combined effect of *DmGS*/*Gclc* RNAi and oxidative stress would be more pronounced than either of the treatments separately. It was found that *DmGS* and/or *Gclc* RNAi combined with a low dose of DEM did have a pronounced effect on cellular viability (Table 7.3). *DmGS* RNAi treatment in cells exposed to DEM did not cause a decrease in GSH levels greater than the DEM treatment (60.5% vs 57%), but the cell viability was significantly reduced compared to DEM treatment alone, and there was a slight, but statistically non-significant, increase in DNA damage. This result is puzzling because it would suggest that GS plays a role in maintaining cellular function that is not directly related to the synthesis of GSH. A similar situation was found by Yang et al. (2002) who studied GCLM null mice, which were viable with no phenotype, however, cultured tissue from these mice exposed to hydrogen peroxide was 10 fold more sensitive than control tissue. GSH depletion was simulated chemically to the same level as in the GCLM null cells, and the knockout cells were still dramatically more sensitive to peroxide killing (Yang et al., 2002). The conclusion the authors reached is also applicable to the findings in this chapter; that cell

death is less dependent on existing cellular GSH than on the capacity for de novo GSH synthesis.

Another point to consider is the GSH assay used in this study. The protocol is based on a glutathione reductase recycling assay (Tietze, 1969) which measures the total amount of GSH by reducing the oxidized form of GSH (GSSG) to GSH. In normal cellular conditions, GSH exists mainly in its reduced form, with only about 1-5% existing as GSSG (Njalsson, 2005), however in the experimental conditions of increased oxidative stress resulting from *DmGS* RNAi and exposure to DEM, it is likely that the amount of GSSG would be significantly increased. Measuring the total GSH therefore did not reflect the probable increase in the GSSG:GSH ratio in cells exposed to *DmGS* RNAi and DEM, which is the ostensible cause of the drop in viability with both treatments.

Similar results were found for cells exposed to DEM and *Gclc* RNAi or a combination of *DmGS* and *Gclc* RNAi (Table 7.3). In both cases cells experienced a significant decrease in GSH and an increase in DNA damage. The increase in DNA damage may have due to the large amounts of ROS released when mitochondria are destroyed in conditions of GSH depletion, as seen in a previous study in mouse embryos (Reliene and Schiestl, 2005). It is interesting to note that in the presence of DEM the viability of cells was not further decreased by the addition of *Gclc* RNAi compared to cells with *DmGS* RNAi alone, even though the effects of *Gclc* RNAi alone or in combination with *DmGS* RNAi are more pronounced with respect to DNA damage and GSH depletion. The lack of a cumulative effect of *DmGS* and *Gclc* double knockdown on cell viability suggests that expression of *DmGS* and *Gclc* are equally important in protecting *Drosophila* S2 cells from DEM-induced oxidative stress, even though their effect on cellular GSH levels is not the same. These results also reiterate the findings of previous studies in whole organisms, as well as cell culture, showing that a reduction in antioxidant gene expression is only detrimental to viability under conditions of

oxidative stress (Mackay and Bewley, 1989; Seto et al., 1990; Radyuk et al., 2003; Van Remmen et al., 2004).

The effect of *DmGS* and/or *Gclc* RNAi treatment in combination with the DEM-induced oxidative stress was not corroborated by exposure of cells to H_2O_2 (Table 7.3). The addition of H_2O_2 to cells which under-expressed *DmGS*, *Gclc* or a combination of the two genes had no further effect on cell viability, which is interesting because GSH is required to protect cells against the $O_2^{\bullet-}$ radical formed by peroxide, as illustrated in Figure 7.7 (Eklow et al., 1984; Dringen et al., 1999). This unforeseen result may reflect the sufficiency of catalase, which is the primary cellular defense against peroxide (Mackay and Bewley, 1989), or it may be due to the different mechanisms of action of each of the oxidative stressors. The manner in which DEM causes oxidative stress is by irreversibly binding GSH, thereby causing the depletion of a key cellular antioxidant (Back et al., 1998). H_2O_2, on the other hand, is a powerful oxidant which causes a rapid oxidation of GSH via glutathione peroxidase, but it does not remove GSH permanently from the system (Eklow et al., 1984; Dringen et al., 1999). The use of DEM as a source of oxidative stress may have been more effective in reducing the viability of cells under-expressing the GSH biosynthesis enzymes *DmGS* and *Gclc* because of the already reduced cellular GSH levels.

The second aspect considered in this chapter was how the level of *DmGS* and *Gclc* transcription affected each other's expression. As discussed earlier, a low dose of DEM had the effect of slightly up-regulating *DmGS* and *Gclc* expression (not statistically significant), while H_2O_2 had the effect of down-regulating *DmGS* and *Gclc* expression. *Gclc* RNAi slightly reduced *DmGS* expression, however, *DmGS* RNAi had the effect of up-regulating *Gclc* expression (Table 7.3). The more pronounced effect of *DmGS* RNAi on *Gclc* expression in the presence of DEM may be a result of the DEM-induced GSH depletion which caused a further increase in *Gclc* transcription by increasing the oxidative stress. These results suggest that

DmGS and *Gclc* are regulated in a similar manner, although the downregulation of *DmGS* has a greater effect on *Gclc* expression than in the reciprocal situation. This could be because *Gclc* is the rate-limiting step, and is more tightly regulated with respect to changes in the cellular environment that affect GSH levels, including *DmGS* transcript levels.

8 Chapter 8: General Discussion

Glutathione depletion has been linked to a variety of human diseases including Parkinson's disease, Alzheimer's, liver damage, amyotrophic lateral sclerosis and cystic fibrosis, in addition to ageing (Anderson, 1998; Schulz et al., 2000; Townsend et al., 2003). The experimental depletion of glutathione by both chemical and genetic means results in altered redox potential, and triggers an increase of ROS, mitochondrial permeability, apoptotic signalling events, and eventually cell death (Meister, 1995; Armstrong and Jones, 2002; Njalsson, 2005; Shen et al., 2005; Hong et al., 2006; Valverde et al., 2006). GSH concentrations have been shown to decline dramatically with age in *Drosophila*, and it has been suggested that the GSH-dependent cellular redox state is a determinant of the physiological age of flies, and may be involved in determining the rate of ageing (Rebrin et al., 2004). Ageing rats also exhibit lower levels of GSH (Zhu et al., 2006), and in mice there is a pro-oxidising shift in the glutathione redox state with age (Rebrin et al., 2003).

The free radical theory of ageing (Harman, 1956) proposes that ageing is caused largely by ROS damage to macromolecules. It has also been suggested that ROS have an effect on the ageing process by interfering with redox gene expression and intracellular signalling pathways. For example Sir2, implicated in CR-mediated lifespan extension,in *Drosophila* is regulated by the cellular redox status (Dugan and Quick, 2005; Guarente and Picard, 2005). The proposed link between oxidative stress and longevity has been strengthened firstly by the identification, in several species, of mutants that have increased longevity and increased resistance to oxidative stress, and secondly by transgenic organisms with enhanced oxidative stress resistance that were also found to have extended lifespans (Parkes et al., 1998; Orr et al., 2003; Morrow et al., 2004; Fridell et al., 2005). Thirdly, the increase in oxidative damage to macromolecules has been observed with increasing age (Yu et al., 1982; Sohal et al., 1995;

Finkel and Holbrook, 2000; Stadtman, 2002; Landis and Tower, 2005). One of the tenets of the free radical theory of ageing is that an increase in antioxidants will result in increased oxidative stress resistance, and ultimately increased longevity. In this thesis, it was proposed that altering the level of GSH would have an impact on oxidative stress resistance.

GSH is synthesised in a two step reaction involving two distinct enzymes. Glutathione synthetase (GS) has received less attention in the literature because it is widely accepted that glutamate cysteine ligase (GCL) catalyses the rate-limiting step in GSH synthesis, although this is based on kinetic data obtained from enzyme studies in the rat kidney (Yang et al., 2002a). A recent collection of studies in mice, humans and *E. coli* has, however, suggested a greater role for GS in determining the rate of GSH synthesis than previously thought (Choi et al., 2000; Huang et al., 2000; Yang et al., 2002a; Njalsson and Norgren, 2005; Liao et al., 2006). This thesis aimed to characterise the gene encoding GS in *Drosophila melanogaster (DmGS)*, and to determine whether altering *DmGS* expression would alter oxidative stress resistance in this organism.

The first objective of the thesis was to determine the structure of the *DmGS* gene. It was found that *DmGS* exists as a tandem duplication in the majority of *D. melanogaster* laboratory strains tested, as well as in all the *D. melanogaster* wild type strains tested. The lack of the *GS* duplication in any sibling species of *Drosophila* tested, and the lack of sequence divergence between the two copies suggests a very recent duplication event. *DmGS* is an 8kb gene present on the X-chromosome, and consists of nine exons which are alternatively spliced, resulting in nine predicted *DmGS* transcripts.

The second objective of this study was to analyse the multiple *DmGS* transcripts predicted from the published *D. melanogaster* sequence. Six of the predicted transcripts were found *in vivo* by cDNA cloning and sequencing, although it appears that only two of the six transcripts are likely to encode an active GS protein. The two transcripts (T4 and T8) encode the same

protein sequence, and differ only in their transcription start sites. Multiple GS transcripts have been observed in mice, which express six separate transcripts that encode three different proteins, although only one of those proteins is active (Shi et al., 1996). The active *Drosophila* GS protein is predicted to consist of 491amino acid residues, which is comparable to the GS protein sizes in eukaryotic species studied to date (Huang et al., 1995; Shi et al., 1996; Grant et al., 1997; Wang and Oliver, 1997; Phlippen et al., 2003; Njalsson et al., 2004). The predicted active *Drosophila* GS protein also shares strong homology with other eukaryotic GS protein sequences, and 21 of the 24 active residues mapped on human GS are identical to the *Drosophila* protein (Njalsson et al., 2004). The transcripts encoding the active GS protein also appear to be significantly more abundant than the larger sized T1 and T2 transcripts in all developmental stages of *Drosophila.* It is possible that the transcripts which do not encode active proteins are involved in GS regulation, although the mechanisms of this predicted regulation are unknown at this stage. Another question which remains unanswered is: which transcripts originate from each of the duplicated gene copies? The two gene sequences are identical, and no indication is given on the annotated sequence as to how the predicted transcripts are assigned to each gene.

It is unclear from this study what effect the presence of the second copy of *DmGS* has on the strains which carry the duplication, since the GSH levels were unaffected by GS copy number, and there was no difference in oxidative stress resistance between single and double copy number strains. The third objective of this study was to determine the role of *GS* in maintaining GSH levels, and in oxidative stress resistance in *Drosophila.* DEM-mediated GSH depletion, which has been shown to cause an increase in ROS generation after 12 hours of treatment (Back et al., 1998), was used as a source of oxidative stress in adult flies as well as in S2 cells. Expression of *DmGS* as well as *Gclc* was found to increase in response to DEM and Paraquat-mediated oxidative stress, both in adult flies, and in the S2 cell culture system, as

supported by previous reports (Ruth Akhtar, pers comm.; Orr et al, 2005). A 58% increase in GSH in response to a low dose of DEM was also observed in cultured *Drosophila* cells.

An EP element inserted in the 5' end of the distal *DmGS* copy was used to drive expression of *DmGS* in combination with a ubiquitously expressed GAL4 driver. *Gclc* was also overexpressed using the GAL4-UAS system, in combination with *DmGS* or separately, to determine the effect of each gene on the GSH biosynthesis rate. *DmGS* overexpression resulted in a lower GSH increase (49.5-63.2%) than *Gclc* overexpression (142-243%) in both male and female flies compared to their non-driven controls, which is consistent with *Gclc* being the rate limiting step in GSH biosynthesis. The combined overexpression of *DmGS* and *Gclc* caused a further increase GSH levels compared to *Gclc* expression alone, highlighting the importance of *GS* in GSH synthesis, although this effect was observed only in female flies.

The rationale for testing the oxidative stress resistance of flies overexpressing GSH biosynthesis genes originates from the idea that laboratory conditions are generally stress-free and do not simulate the variable and often stressful environmental conditions to which wild type animals are exposed. Overexpression of antioxidants would be expected to have a greater effect on lifespan when assayed in conditions of severe oxidative stress (Lin et al., 1998), compared to unstressed conditions. This was corroborated by the evidence in this thesis which showed that female flies overexpressing *DmGS* and/or *Gclc* had statistically significant increases in stress resistance. A recent study overexpressing GCL in *Drosophila* showed a 50% increase in lifespan, and higher resistance of flies to oxidative stress (Orr et al., 2005). The increase in both median and mean survival implied that there would be a delayed onset of mortality rate increase, as well as the overall length of lifespan, through *DmGS/Gclc* overexpression.

An interesting sex difference was noted in flies overexpressing *DmGS*: males had higher *DmGS* transcript levels, as well as higher GSH concentration than their female counterparts, although an increase in oxidative stress resistance in *DmGS* overexpressing flies was only observed in the female flies. This result was counterintuitive, and may be an artefact of the gender differences which have confounded many previous studies on oxidative stress and ageing in *Drosophila* (Lints et al., 1983; Carey et al., 1995; Spencer et al., 2003; Forbes et al., 2004; Magwere et al., 2004; Valenzuela et al., 2004; Magwere et al., 2006). As yet, the reason for the sex biases often observed in lifespan studies is unknown, although there are theories suggesting reproductive cost, sex linkage of genes, and difference in male and female metabolic rates (Vina et al., 2003; Burger and Promislow, 2004).

The effect *DmGS* has on GSH levels, as well as oxidative stress resistance was corroborated in the *Drosophila* S2 cell culture system using RNAi, thus fulfilling the final objective of this thesis. Although RNAi-mediated *DmGS* knockdown alone did not affect GSH levels, *DmGS* RNAi did have a significant cumulative effect on GSH reduction in the presence of *Gclc* RNAi. Because RNAi results in a knockdown of the gene, and does not completely obliterate mRNA molecules, a low baseline of activity of GS or GCLC remains despite treatment of cells with dsRNA complementary to a portion of the *DmGS* or *Gclc* genes. This may explain the lack of effect of GS/Gclc RNAi on cell viability and DNA damage. The ability of cells to survive with very low levels of GSH (as low as 2% of wild type) has previously been observed in a murine *GCL* knockout system, although the cells did 'teeter on the edge of survival' (Valverde et al., 2006). A complete withdrawal of GSH led to initiation of apoptosis after 24 hours, although the cells continued to synthesise DNA and divide for approximately 48 hours in the absence of GSH (Valverde et al., 2006). These findings may explain why no effect on viability was seen with *DmGS* and *Gclc* knockdown, since the maximum reduction of GSH was ~50%, and the time period of the viability assay was 48 hours.

As was the case in adult flies, a change in the expression of GSH biosynthesis genes had the greatest effect on viability when S2 cells were exposed to an oxidative stressor. *DmGS* or *Gclc* RNAi halved S2 cell viability in the presence of a low dose of DEM. This effect on viability is not solely due to extreme GSH depletion in the presence of DEM, as shown by the similar GSH levels in cells treated with DEM, with or without *DmGS* RNAi, although the significant reduction in viability was observed only with the combination of DEM and *DmGS* RNAi exposure. This result is supported by evidence from a study on *Gclm* null mice which exhibited no phenotype, although cultured cells were 10 fold more sensitive to the presence of peroxide stress (Yang et al., 2002b) compared to control cells. These results together suggest that cell survival is dependent on the capacity for de novo GSH synthesis in the presence of an oxidative challenge, rather than on the amount on GSH present in the cells. The ability of *Drosophila* to survive in conditions of oxidative stress with reduced antioxidant levels was observed in SOD hypomorphs (Seto et al., 1990), as well as in catalase hypomorphs (Mackay and Bewley, 1989).

The combination of DEM and *Gclc* RNAi caused a reduction in viability, as well as a significant increase in DNA damage. Although *Gclc* knockdown causes a greater incidence of DNA damage than *DmGS* knockdown, there was no cumulative effect on cell death when both *DmGS* and *Gclc* expression were reduced. It is possible therefore that GS and Gclc are equally important in protecting S2 cells from DEM-induced oxidative stress, despite different effects on GSH concentrations. Once again, this draws attention to the idea that altering the levels of antioxidant genes has an effect only under oxidative stress conditions, as seen in previous studies using *Drosophila* adult organisms or cell culture (Mackay and Bewley, 1989; Seto et al., 1990; Radyuk et al., 2003; Van Remmen et al., 2004).

Overexpression of 'major' cellular antioxidants, such as MnSOD, catalase and Gclc in ubiquitous tissue distribution has been shown have no effect, or to be detrimental to

Drosophila viability (Bayne et al., 2005; Orr et al., 2005), probably due to the damaging effects of extreme ROS depletion on cell signalling. Conversely, overexpression of major antioxidants specifically in neural tissues, has a positive effect on oxidative stress resistance and longevity (Parkes et al., 1998; Sohal et al., 2002; Landis and Tower, 2005). In relation to these findings, this thesis suggests that the ubiquitous overexpression of genes with a more minor role in antioxidant protection, such as *Gclm* (Orr et al, 2005) or *GS* (this thesis) can have a beneficial effect on *Drosophila* viability under conditions of raised oxidative stress. The difference in ubiquitous expression between major and minor antioxidants is probably a result of the minor antioxidants not decreasing ROS below the baseline value where the reduction of ROS has detrimental effects by disrupting signal transduction pathways and gene regulation (Phillips et al., 2000; Mockett et al., 2003; Bayne et al., 2005). Under normal conditions there are probably optimal levels of *DmGS* present, however under conditions of oxidative stress, an increased level of *DmGS* is beneficial.

This thesis could also be seen to support the free radical theory of ageing. Since the overexpression of *DmGS* resulted in increased survival under conditions of oxidative stress in female adults, and the knockdown of *DmGS* in S2 cells resulted in a decrease in oxidative stress resistance, the expectation is that flies overexpressing *DmGS* would also have increased lifespan, although these experiments could not be performed due to time limitations. These two findings highlight the negative effect of *DmGS* knockdown on cell viability only in conditions of increased oxidative stress emphasises the idea that there is a degree of redundancy in the cellular antioxidant defence, and when ROS production is increased under stressful conditions, the upregulation of the antioxidant system would be beneficial for the rapid removal of the toxic species. Under normal physiological conditions when ROS production rates are relatively slow, however, the increase in antioxidants may not have any additional beneficial effect.

The overexpression of *GS* would therefore be expected have a positive effect on oxidative stress resistance and possibly protect against early death in non-laboratory organisms which are exposed to stress, pathogens and toxins on a daily basis (Dugan and Quick, 2005). This finding, as well as the positive effect of ubiquitous *Gclm* overexpression encourages further work on minor antioxidants which may reduce ROS to a level where oxidative damage is decreased while maintaining functional cell signalling pathways and gene regulation.

8.1 Future work

Further work that could arise from this thesis includes lifespan studies of flies overexpressing *DmGS* alone, and in combination with *Gclc*. In the S2 cell culture system, *DmGS* could be overexpressed using the cDNA of the transcript predicted to produce the active GS protein (T4/T8). This could confirm the results of the *DmGS* overexpression experiment performed in adult flies, and could also determine whether there is a GSH level above which there is no beneficial effect on oxidative stress resistance. Further RNAi experiments could also be performed in the S2 cell culture system. It would be interesting to investigate the effect of knocking down the expression of specific *DmGS* exons compared to the total *DmGS* knockdown performed in this thesis using dsRNA specific to unique exons. This would help determine the function of the different transcripts, and identify whether all transcripts play a role in cell survival. A similar approach has been used to study the 95 alternative exons of *Dscam*, where dsRNA fragments were created to specific exons to facilitate RNAi in S2 cells (Schmucker et al., 2000). A further technique to determine the location of specific transcripts within larvae or adult flies would be the use of in-situ hybridisation. Probes with sequences identical to unique portions of each of the transcripts would bind to the tissues or cells in which the corresponding transcripts are present.

RNAi-mediated knockdown of total *DmGS*, as well as the specific transcripts, could also be performed in adult flies using the pWIZ plasmid in which a transgene is inserted in two orientations separated by an intronic sequence, resulting in the production of dsRNA with a hairpin structure (Lee and Carthew, 2003). This endogenous RNAi system has the potential to be stably inherited and can be controlled by the GAL4/UAS system. Adult flies with *DmGS* knockdown could be exposed to oxidative stressors to confirm the results obtained using the S2 cell culture system in this thesis, and the effects of *DmGS* knockdown on *Drosophila* development could be observed.

A technical addition to this thesis could be including GSSG measurements with the GSH measurements. The Tietze method used in this study measures total glutathione levels, including the reduced and oxidised forms. With access to an HPLC facility, accurate measurements of the oxidized form of GSH could be taken, and may shed light specifically on the experiment where S2 cells were treated with *DmGS* RNAi and DEM, in which it is likely that the amount of GSSG would be significantly increased. It has previously been noted that upon exposure to an oxidative stress, the oxidation of GSH to GSSH, coupled with rapid GSH synthesis can account for the lack of a decrease in total GSH levels (Coe et al., 2002). The ratio of GSH:GSSG provides more information in oxidative stress studies than total GSH alone.

9 Appendixes

9.1 Appendix A: Region of BACR29H04 containing both copies of GS (position 22880 to 42300).

The P-element insertion point in strain EP1322 is highlighted at position 25010
The 297 element is highlighted (position 28890-31619), and is present only in the distal copy of *DmGS*.
The mRNA of the distal *DmGS* copy starts at position 24784-33692, and the proximal *DmGS* copy mRNA is at position 34939-41115.
The duplicated region spans positions 23700-41280, with the duplication starting at position 33860.

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      22890      22900      22910      22920      22930      22940      22950      22960
CTTTACTTTA AATTAATTTC CGCTCTATGC ATAACCAATG CATCGCATTG GGACCACTAG TTCACGACAG CTTGAGTTGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      22970      22980      22990      23000      23010      23020      23030      23040
ATTTTTTAAT GGACTTTTTT TTTTTTTTTA ATTCAAATCC TTTGCAATTA AAAATTTCAT TTCCATTTAA ATGTAACTTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23050      23060      23070      23080      23090      23100      23110      23120
CAACCGGTTT TAATTAGAAC ACCCAAATAC AGTTATTTTG AAAAACGAAG AAGTTGTAAA ATTGTGACCC GATTTAAAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23130      23140      23150      23160      23170      23180      23190      23200
GTTAAGCATC ATAAAAGTGA GGCACTGCGA GAATATAATG ATTTGTAAAT ACGATTTTTG CATTCACATT ATATATTATA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23210      23220      23230      23240      23250      23260      23270      23280
AAATCGAATT TATTTAAAAA GAATGCCCAT TTATCGCTGT GCAAAGTTCA AACCCATTTC ATCCTCTGTG CGAATTTATG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23290      23300      23310      23320      23330      23340      23350      23360
TGGACACTTA AACTCGCGAC TTAAAATGCC GAAGACTCTA AAGCAACACG TTATTAAAAA TTTGATGGGC ACATCAAGTG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23370      23380      23390      23400      23410      23420      23430      23440
CTTATAAATA TTCATTAGCG CATAAAAAAT CAGTTTATTT CGGCGCAGCA ACTGGCACAG TTCGCATGTT CAGGATAATG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23450      23460      23470      23480      23490      23500      23510      23520
AAAACCGCAT ACCTCATCCT GATTCTTTCG CTCTCGATAT CCACTTCCTT GGCTAACGAT CCGCAGAGGT CCTTCGATCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23530      23540      23550      23560      23570      23580      23590      23600
GAATGCGGAC AATCCGTGCA GTGCTCACCG CATTTGGGTT TGGAGACCCA TCCTCGATCG ATTCGCCAAG GTGAACTCCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23610      23620      23630      23640      23650      23660      23670      23680
GCGATGAGCT GAAGAACCAG GTCAACGCCA TGTCGGAAAC CATCAAGGAC CTGCAGACGC GCATTGGCCA CAAGAACGAG
                Start of duplicated region
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23690      23700      23710      23720      23730      23740      23750      23760
CTGCTCCTGC TGCAAGCCGG TCAGATCAAG GACCAAAGTG CCAGGCTCAA CGTTTTGATC AAAAACATGA AAGTCCTTCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23770      23780      23790      23800      23810      23820      23830      23840
AAGTGAGTGT CTGAATGCCC AGGCGGAGAT CAAAAGCAAG GACATACAAA TACAATTGAG CGCGGCAAAG ATCAAGCAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23850      23860      23870      23880      23890      23900      23910      23920
```

```
TGAGCAACGA ACTGGTGACG CAGTGCAGCC GCCAGGACAC GTGTCCCATC GACGGCAAAG GTGGCATCTA TAAGCTGAAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      23930      23940      23950      23960      23970      23980      23990      24000
ATCCGTGAAT TGCCAGCATT CGAGGCGCCG TGCAGTTCCA ACGGCTGGTT GACCATACAG AAACGATACG ATGGCGCCGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24010      24020      24030      24040      24050      24060      24070      24080
GAACTTCGAT CGCGGTTGGA AGGACTACAA GGATGGTTTC GGACGCGTGA GAGGTGAGTT TTTCATCGGC CTGGAGAAAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24090      24100      24110      24120      24130      24140      24150      24160
TGCACCTTAT GACGCGGCAA CGACGCCACG AACTGTACAT TAAGCTGGGC AAGATTGATG GCACCACTAG CCACGCCCAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24170      24180      24190      24200      24210      24220      24230      24240
TACGACAACT TCGAACTCGG CGGCGAGATT GAATCCTATG AGCTAAAGTC ACTGGGCCGC TACAATGGCA CAGCCGGCGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24250      24260      24270      24280      24290      24300      24310      24320
TTCTCTGCGG CCACATGAGC GCCAAAAGTT CACCACCAAC GACAAGGACA ACGATGCCTA TCGGTTCAAC TGTGCCGCCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24330      24340      24350      24360      24370      24380      24390      24400
ACGAATATGG TGGCTGGTGG TACTACGATT GCGCCAAGAG GTAAGTACGC CCATTATATA CATATGGTAT ATATACATAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24410      24420      24430      24440      24450      24460      24470      24480
GTACTTGCAT GTAATATTAA TCGGCGGAGT AACTGCAGCA TGCTCAATGG CAAGTTCTAC AAGGAGGGCC GCTCGCGGAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24490      24500      24510      24520      24530      24540      24550      24560
TGGAAAAACC AATGGCATCC TGTGGGGCTC CTGGCACAAC AACGATTGGA CCTACTCGCT AACCTTCGTG GAGATGATGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24570      24580      24590      24600      24610      24620      24630      24640
TAAGGCCGCG AATAGATTAG CTATTACAGT AAGAATAAAA AAATGGCACA AATAAACGAA AAGGTCTCAT TTGTAATCAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24650      24660      24670      24680      24690      24700      24710      24720
TACATTCAAT TTCAATTAAA TTAAATAACG GGCACTTACT AAACTTTTCT CGCCGTGTAT GCCCCTGGTA TTTTTTGATG
                                                                  Start of distal DmGS copy
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24730      24740      24750      24760      24770      24780      24790      24800
CCATATTCTG GCATATTGCG GTATATCGCC GGCTGGTATT TTTTGCCTAA CCCGCTACGC ACACACTGCG AATATTCGGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24810      24820      24830      24840      24850      24860      24870      24880
CAAAAACATA GTTATATTTT TGTCTGTGGG CTTTTAGACA AATTGTAAAC TGATAGCTCC CGACCCCAGC CGTTTAACGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24890      24900      24910      24920      24930      24940      24950      24960
TGTCCAAATT GACAGTGCAA CCGGTGGAAC AGGTGGAGCA CAAGCAGCGG CACCCAGAAG GAAACCAGCG ATCCAGTCAC
                                                    P-element insertion site
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      24970      24980      24990      25000      25010      25020      25030      25040
AACAATCGCT CGTTGAGAGC GCAGCGATAA GAGTTCCAGA GTTTCTAGAG TTTCCAGCGA TACCCAGCAG TGGTGAGTAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25050      25060      25070      25080      25090      25100      25110      25120
CTAGTAGACA GTCGTCGCAC TGCTCAGCAG TGCTCAGTGC CCAGTGAGCA GTTCCGGCGA GCAGACCCCG TGAGATACCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25130      25140      25150      25160      25170      25180      25190      25200
CGGCGAACAA CTAACAACTA AGATACTCTG CCCCTTTTCT AGACGCAGCT GAGCCCCAGA TCCCAGATCC CAGATACCAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25210      25220      25230      25240      25250      25260      25270      25280
CAGCAGACCG ACCCAAGCAG AACGATCATG TCCAGCGACG CCAATACGGC CGTGCTGCGC AACTGCATCC GCTTGCCGCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
```

```
      25290       25300       25310       25320       25330       25340       25350       25360
GGCGGAGGAT GAGCTACTGG AGGTGACGGC CAAGGCCAAG GATTATGCCA TTATGCACGG CGCCGCCATG CGATCGAAGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25370       25380       25390       25400       25410       25420       25430       25440
CGGCCTTCAG TCCGGACTCG TTGAATGTGG GTGATACAAT GGGCACTCGC TAGTGAGGGC TCGTATTCAT GCTAAGGTGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25450       25460       25470       25480       25490       25500       25510       25520
CTAGTTTCCG CTCAAGAAGC TCATTTCAAT GGAACAACAT TCATTACACA TCCCTTATCG AGTTCAGTTC CTTTTTAAGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25530       25540       25550       25560       25570       25580       25590       25600
TATGTGTAAT GCCAAATGTG ATTGGATGCT TGAGTTTGTT TTTCATTTTA TTGACGGTAG AGAAAGATTA GCAATGTGAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25610       25620       25630       25640       25650       25660       25670       25680
GGTATAAGCT ACACATTGGC ATTCAATCGA GCTTCAAGCT AGCGTGCTGC TGCTTGAGCA ACAGTTTGCC AGGATACTAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25690       25700       25710       25720       25730       25740       25750       25760
ACATTTTCAT TTTCCAGTTC GCCCCCTTCG TGCTGGTGCC TTCCTCGTTT CCGCGCAAGG AGTTCGAAAA GGCGGTGGCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25770       25780       25790       25800       25810       25820       25830       25840
CTACAGCCGA TCATCAACCG GCTGATGCAC AATGTGGCCC ACGACGAGGA GTTCATCACG ACGACGCTGG CGGAGACGAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25850       25860       25870       25880       25890       25900       25910       25920
CAAAGTGGAC GAGTTCACGG CCAATCTGTT CAACATCTAT CGCAAGGTGC TGGCGCACGG ATTCACCCAG GTGAGATTCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      25930       25940       25950       25960       25970       25980       25990       26000
GTCAGCTGGC GTCCATCCAT CTCCGATCCT CCGCCCTCTG CACACCATCA TCGTATCACC CGCCAACGTA TCCACCTCAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26010       26020       26030       26040       26050       26060       26070       26080
TCCCATCGGG AGGGAGCACA TCTACAGGCA GCCAGTGGGT TACTCGCTAG AATTTATTGG GCCCCTCCAC CCGCGGTCGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26090       26100       26110       26120       26130       26140       26150       26160
AATCTAATCG CCATTCCATA AATCGCCCGG CCATAGCCAT AGCCGCAGTC AAAGCCCAGA AGTCGGCTGT AGTCGCTAAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26170       26180       26190       26200       26210       26220       26230       26240
CTCGCGTGAT TTATGCGAGC ATCCTGGCCT CACAGCTCCT TATCGTTGGC TTGTAAATCG ACGTAGGACA GCCCATCCCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26250       26260       26270       26280       26290       26300       26310       26320
GACCCCCATA ACGCTAGAAC TTTGCCATTG CTCCTCCTAC TTTGATTTAC ATGCATTTCC CCTAACGAAA CCAGCACAGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26330       26340       26350       26360       26370       26380       26390       26400
GGCGAATCCA TTGAGCTCTT CCACGTCCAC AAGAAGCTTT AATCTTTGTA GATCCTATCA GAAAGGTGAC AGTTTTACAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26410       26420       26430       26440       26450       26460       26470       26480
TTTGAATAGT TATTGGCCCG GTAACTGATA GGGTTTATTA AATATAACAA TGCCAGCGCT ATTTTTACTT TATTACTATT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26490       26500       26510       26520       26530       26540       26550       26560
CAAACTTTGT CTCGAATCTT ATCGAAGTCT CAATTGAAGC ATTTAAATGC TAGAATCAAT TAAACTATTG TTCATTTTAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26570       26580       26590       26600       26610       26620       26630       26640
ACTCGTTGAT TTTTGCCAAT GATTTTCAAT GTTTTTACTT GCAACTTAAT CATTTTCATC ACCCAATATC AAGTTTAGTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26650       26660       26670       26680       26690       26700       26710       26720
GCGAATTTTG CTCTCTGTGC AAAAATTTAG TTCACGATCT GTGCTTTTCT CGTCACCACT CCTTCAAAAA TTTTAAAACC
```

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26730      26740      26750      26760      26770      26780      26790      26800
TAACCACGCA ATTGTACAAA CTCTCGATCT CAAACGCAAT CTGTTGCTCC ATTATCCGTC TCTCATTGCA TCGCTGTGCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26810      26820      26830      26840      26850      26860      26870      26880
ACGAAAAATG ATTTTGAATC AAAGAAAACA TCGCTGGGGA TGCTGCGCAG CGATCTGATG CTGGAGTCCG GCTGTCCCGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26890      26900      26910      26920      26930      26940      26950      26960
GCTGTCGCCA CGGGCGCTGC GAACGGCGGC CGGTGAGGAT CGTGGTCAGG ATGCGGGCGC TGCTGTGGGG CAAATAGCTG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      26970      26980      26990      27000      27010      27020      27030      27040
GAGCTAACGG AGCTGCAGGA GTCGGAACTG CCGCTGGCAC GGGCAGCAAG GAGGAGGAGC ACAGGGAGGT CCAGCTAAGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27050      27060      27070      27080      27090      27100      27110      27120
CGGGTCACCA AAGAGCCGGA ACGGAGGGCG ACCGGCGCCC AGTCCGCTTA CTGCTGCTGG AAGCAAGTCG AGATCAATAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27130      27140      27150      27160      27170      27180      27190      27200
GATAGCATCG GGCTTTGGCC ACTTGGGACC AGCAAGTAAA ACCATACAAA GGTGCTGTTC ACGCGCACAG TGCATCCATC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27210      27220      27230      27240      27250      27260      27270      27280
AGTGCAACTC CGCACCACGC CATCATCAGC AACACACCCC CTAAAAAAAA AAACCCAGTA TTCCTTCCAT GACACGTCAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27290      27300      27310      27320      27330      27340      27350      27360
TCATCTCATC TCATCGCATT TTTGCATTCC TGCATCGCTT GCATCCTCCG CCTAACCGCA CAAAATCACA CACTAACTAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27370      27380      27390      27400      27410      27420      27430      27440
TTTCGATACC CGAATCAAAT ATTAACTATA TTTGAATCCT TAACTTTGTC AATCGAAAGA ATCTTTTGCA TCTGAAAAGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27450      27460      27470      27480      27490      27500      27510      27520
AAGAGCACTA AAGTCAAAGA TAGATCGGAA TTAAAACGAA GTTGCAACAA GAAGTGCATA ATCAGATATT TCGACTGAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27530      27540      27550      27560      27570      27580      27590      27600
TAAGCTATTT AAAATATTTA AAAAATATTT CAAACATATC TTGTAGTGTG TGCTTTTGTG TGTGTTGCTT TGTAAAATCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27610      27620      27630      27640      27650      27660      27670      27680
TTTGTACTTT AGTTGAAGTT GATTTTGTTC CGCTGTTCCT TAGAAATGTA TTTATGATTA ACCATAGCTA CAAACGTTGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27690      27700      27710      27720      27730      27740      27750      27760
TCCTCCAGAT TGTTGCCTAG AGTTTTCATT TGCTGACGCT GCTGTCGCCT GTCGCCCGTT GCCTCCTGCC TGCCTGCTCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27770      27780      27790      27800      27810      27820      27830      27840
GCTCCTGCCT CCTGGCATCC CGCATCCTGC ATCCTGGCAC CTGCCAACCG ACACCTGCTC CTCCCGACCC AACCGTAGCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27850      27860      27870      27880      27890      27900      27910      27920
TCTAACCCGC CCGCTAATAA TCTTCACTTC CGTCAGTGAT TGTCCGTTCT GCTGTCGCTT ACTCTACCAA CTCTATAGTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      27930      27940      27950      27960      27970      27980      27990      28000
GGTCAGGGGT TAGCCAATAT ATATGCATAT ATATCTTTCT AGAAGATTTG AAGATAGCGA TTTCACAAGA TCTGATAAAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28010      28020      28030      28040      28050      28060      28070      28080
TGTTTGGATG ATACATCACA TCAATTTGTC ATTTGTAATC AGCAAAAGGC GCTAATCAAA AATTTCACCA CTTGTACGCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28090      28100      28110      28120      28130      28140      28150      28160
TGGGCTATTA AAACCCTAGA AGTTTCTGGA GGTTGAATTA AACAATTTAA TTGTTTGAAA GAACAAATGT AGTTTTTACT
```

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28170      28180      28190      28200      28210      28220      28230      28240
TGCATAGATA TTAAACTCAC TCAACATTTT GGTCATACGT CGTTTTAGAC GAATCACACC TGTAGTTAGG AAGCCCCACA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28250      28260      28270      28280      28290      28300      28310      28320
TGGAGAGCCC ACAGCTGTAC TGCAGATAGG GTTACAGATA GGGATACGCT GATATAGATA CGCCTGGTAA TTGTTTATTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28330      28340      28350      28360      28370      28380      28390      28400
AAAGTCGAGG CGTGTGCGAT AACAGACCCA ATAAGTGCAT TTGTGGGTGC AAAATGTTTG TCTTCAATTA CAAATTCAAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28410      28420      28430      28440      28450      28460      28470      28480
ACTCGAATTT ATGAAATGCC AGCAGAACTC AAAACACTTT CATGCACGTT TCGAGTTTTT AGCTGTTTGT CTATTTGTTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28490      28500      28510      28520      28530      28540      28550      28560
AATTGCCGGC AATGGTGACG TAAACATAAT TATAAACAGA TTATATCAGA GCATGTTTAT AAAATAGTCA AATTACCAGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28570      28580      28590      28600      28610      28620      28630      28640
TGCAAATGGG CTAATAACTT GTTATTACTT CGTCATGCTT AGTTACATAT GGAAAATAGG TATTAGTAGT AGTATTATTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28650      28660      28670      28680      28690      28700      28710      28720
TTATTATTTC ATTCAATAAC AGCTTGTAGT CATATCGGAA ATATTTAAAC AATTTGAAAC TGGCGGCGTT TGTGGCGAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28730      28740      28750      28760      28770      28780      28790      28800
TATTTGGAGT CTTTCATGGC CCTAAAAATT ATAGAGTTAA TAGCATTGGG AGATGCATCG TTTCTTGCAC CATTTCTCGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28810      28820      28830      28840      28850      28860      28870      28880
CTCTCTCTCT CTCTCCGTCT CTTTTTGAAT ATATATAATT AAAAATTCGC ACCTGCCTCA TCCTACATAT ATATACATAT
       Start of 297 element
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28890      28900      28910      28920      28930      28940      28950      28960
ATATATATAG TAAAATTGTT GTTTTATTTA TGAGTCGAAC TAATGTCCCG TCAGTTGACT AAGGACAACG CCAAGAAATA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      28970      28980      28990      29000      29010      29020      29030      29040
AAAAATAGTT TTTAGGAGAA ATAGTTTCAC GACTTTATTC TTTGGATCTC AGCTTAAGAC TAAAAATCAA ATGCAAATTG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29050      29060      29070      29080      29090      29100      29110      29120
GATTGAGGAG AAGCCTCAGT ATTTGAACAA TATTTGTATT TCGGGAGAGC CTCGCTCTTG GGTTCTTAAG ATTAGGTAGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29130      29140      29150      29160      29170      29180      29190      29200
GTTGAAAGAG GGCAGTCAAA TCTGCTTAGG TCAGGCTTTA GGATGTTTAC AAGAACGGCT GCGCTGCCAA AGATAAACGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29210      29220      29230      29240      29250      29260      29270      29280
ATGTTGCGCA GCTGGGATAC GTTATGGAAA ATTACTAGAG TGCATTACTG ATTTCTTTGA AATAATGGAA GTGGCTGGTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29290      29300      29310      29320      29330      29340      29350      29360
TGGACCACCC AAATACGTCA CTCCCCTTCC CTTAAAGAGA AGCCTATACT TGGGCTGGGA TTGATGGATA TAGTGTGGGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29370      29380      29390      29400      29410      29420      29430      29440
AATGGAGTTT CTTCTATTTC TGTTTGTTGT GGTGTGTTTT GATTTTGATT TTTTCTTATT TTAAGTTTTG CATACAGCAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29450      29460      29470      29480      29490      29500      29510      29520
CATAAATGGT TTAAATGATA CATATCTTAA ATATGAGTAC AACAAAATTA GTGCGATTAT TACAATGATC ATTTGTATGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      29530      29540      29550      29560      29570      29580      29590      29600
```

AAAGTGTAAT ATTGTTCTCT GAAATCATCT CAACAATGTC GTTGTGTTTT ACAATTTTTA TTTTTGTTAT ATTAAGTGGT

29610 29620 29630 29640 29650 29660 29670 29680
GTGTATAGTG GTGTCAAATC TATTTCTGGT TCTAACAGAT TTTCACTTAA ATTATATTAT TGATTTCTAT TTTGCATTGC

29690 29700 29710 29720 29730 29740 29750 29760
GTTACTCTTA TCATTTTGTT TCCTTCTATT TTTATTTTAT TAATTGAATT TTGGCAATTT TGGTTAAGAA TTGTTTGGGT

29770 29780 29790 29800 29810 29820 29830 29840
TAAGTTCCAG GTTACAATTG TGTTTGGTTC TATGTATTTT ATTATTTCGT TCGTGTGTAC TGGCTTAAAA TTACAAATTG

29850 29860 29870 29880 29890 29900 29910 29920
GATTTAAGTG TTTAATAATA TTGGTGACAC ATTCATTTTT TACTTCTTTA TTTTCGGTAT TATAAACTTT ATTTTCTTTT

29930 29940 29950 29960 29970 29980 29990 30000
TCAAAATATG ATTGTGTGTC TGTGTAATCT AGCTGATAGC CGTTGGAGTC TGGGTAAGGA ATTATTTTAA TTGTGCTTAT

30010 30020 30030 30040 30050 30060 30070 30080
TAGTGAAAGG TAAATGGGTA TGTGGGATAT GATTAATACT TGGTTATCAT TGTAGTTAAT CCAAGTGGAT GTTTTAATGT

30090 30100 30110 30120 30130 30140 30150 30160
TCAAAATGTT TTGGCTGTTC ACATTTTCAA GTTTGTCGTA GTTTAGTAAT TTGGGGTTAA ACAGTCCAAG TCTGGAAAGC

30170 30180 30190 30200 30210 30220 30230 30240
TGCATAGCCA TTTCCAAATC TTCTATGTAT TCCGTAAACT GCATTAGTTC GAACAAAAGA CTGTCAATGA TGGTGTGGTT

30250 30260 30270 30280 30290 30300 30310 30320
ATTTTCATAA TTCTGTATTT TTTGCATTCC GTCATTTATC AATCGTATGG CGTCGTTGAG TTCATGTAAT TTTACTGAGT

30330 30340 30350 30360 30370 30380 30390 30400
TATGGACGTT GATTTCAAGT TTTTTCTGTA TTTCGACTCG ATCATTTTCA TCTAGTGTTC CAAATAGGTA TTTAAAACCT

30410 30420 30430 30440 30450 30460 30470 30480
GAACCCACAA TGTTAATAAG ACCCCGTTTG TTTCTATGTT TCAGGGTTAT TCCGGCTAGT TCTCTTTTTA TTTTATTGTA

30490 30500 30510 30520 30530 30540 30550 30560
TAAAAAATTG ATCTGGGGTG CATGGTAATC GGTTCGTAGT CTTTGCTCAA AATAGTCAAC CACGATGTCT ATTTCAGTTA

30570 30580 30590 30600 30610 30620 30630 30640
AATTGATTTT TAAGGAGTGA TGTTCAAAGG AGGATGGTAT CTGTACTGGC TTATCAGAAA AAAGGAGATA TCCGTGGTTG

30650 30660 30670 30680 30690 30700 30710 30720
GTGTCAATAT TATTAATTTG AATTTGTTGT CCATGAACAG CTGTGATCAG TATAAACAAT AGTGTTATCG GGTACCAGGT

30730 30740 30750 30760 30770 30780 30790 30800
GCCTGTGAAA TTGAGAGTTT ATTATTTTTC TTTTGTTTTT TGAATTGTGT TTTGTAGTAG TGCGTAACTC TATTTCTATT

30810 30820 30830 30840 30850 30860 30870 30880
AGTGATTTTG TAATGGTCAG GGTCTGTTTG TTCTACATTT TTGGTTGGTT TAAATGGGTT TTCTAATTTG CCTTTCTGTA

30890 30900 30910 30920 30930 30940 30950 30960
GTGGACCTTT TCTGTAATTA GTGTCAATAT TAAATTCCTG TCTGTCTTCA TTTATTTTGT CTATTTTCTT CTCTTTAATT

```
      30970      30980      30990      31000      31010      31020      31030      31040
TTTTGAGTGT CTAATATGGG ATGCCCAGCG TATAAGAAAA TTTGAGCAGG TGTCTGTCCA GTAGTGTCAT GTTTAATTTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31050      31060      31070      31080      31090      31100      31110      31120
ATGGTTGTAT GTGTAGAGGA TTGTTTCTAT CTTGCTTAAT TTTACTTCTT CATCATCAGA TGAATTGATT ATACGAATTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31130      31140      31150      31160      31170      31180      31190      31200
TTTCATTTAT TGTTTTGTGT AATCTTTCGA CGTCTGCTAC CCCGTTTTTT GCTGTATTGA GCTGTAATCC TACCGACTGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31210      31220      31230      31240      31250      31260      31270      31280
GCCAGTAAAA TTGTTGTTTT ATTTATGAGT CGAACTAATG TCCCGTCAGT TGACTAAGGA CAACGCCAAG AAATAAAAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31290      31300      31310      31320      31330      31340      31350      31360
TAGTTTTTAG GAGAAATAGT TTCACGACTT TATTCTTTGG ATCTCAGCTT AAGACTAAAA ATCAAATGCA AATTGGATTG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31370      31380      31390      31400      31410      31420      31430      31440
AGGAGAAGCC TCAGTATTTG AACAATATTT GTATTTCGGG AGAGCCTCGC TCTTGGGTTC TTAAGATTAG GTAGCGTTGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31450      31460      31470      31480      31490      31500      31510      31520
AAGAGGGCAG TCAAATCTGC TTAGGTCAGG CTTTAGGATG TTTACAAGAA CGGCTGCGCT GCCAAAGATA AACGAATGTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31530      31540      31550      31560      31570      31580      31590      31600
GCGCAGCTGG GATACGTTAT GGAAAATTAC TAGAGTGCAT TACTGATTTC TTTGAAATAA TGGAAGTGGC TGGTTTGGAC
                End of 297 element
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31610      31620      31630      31640      31650      31660      31670      31680
CACCCAAATA CGTCACTATA TATATATATA TATGCACGTA TATATCTATG GACGAGGATC GTAAACCATA CCTAGACATA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31690      31700      31710      31720      31730      31740      31750      31760
TTGAATCGCT TTACCGGCTA AGCAGCTAAT CTTTAAGGGA CGCACCTAAT GTATATATAT CTTGTATAAC GATCATAATT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31770      31780      31790      31800      31810      31820      31830      31840
GCAGGCATAT AGCATCGGTC TGTTGCGATC CGACTATATG GCACACGTGT CCGAGGGCTG TGCGATAAAG CAGGTCGAGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31850      31860      31870      31880      31890      31900      31910      31920
TCAACACGGT GGCCTCCAGT TTCGCTGGCA TAGCCACCCA ATTGGTGCCG CAGCAGAAGT ATTAAGCGCT CTTCTCTCTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      31930      31940      31950      31960      31970      31980      31990      32000
TCTCTTTCTA CATATAATAT ATATAAATAT ATATCTATCA AACTACATAT ATTCCTTTTA CCAAACGCAC GAATGCACAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32010      32020      32030      32040      32050      32060      32070      32080
TGCACACACC TAAAAGCAAA AGCACACTAC ACACCTGTCC TAATATCCTT CCACCCAACT CCTTCTCCTT CTCACTGATC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32090      32100      32110      32120      32130      32140      32150      32160
ACTTATCCAC CTACAGAGGT GCCATCTCTA TCGTTTGTTG TCTCTTTCTG CATGTGTTTT TTTTTTTTTT AATTGGTTAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32170      32180      32190      32200      32210      32220      32230      32240
GGCTCTATCC GTCTCTTTCC ACACGCTGTA GCTTCATAAC CCCACTTCCA CAATTTCCCT AACCATAAAC ATACATATGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32250      32260      32270      32280      32290      32300      32310      32320
ATATCCATCC GCCCATTTGA GAGGTTTTGG TCTCAGTAAT TAACCCAAAT TTCAACCCCC CCCCCCATTA CCACCGAACA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32330      32340      32350      32360      32370      32380      32390      32400
TTTGTTGTTG TTTGTTGCTG TTGTTGGTCG CTGAAAGACT GCGGAATTAT GCTAATGTCT GTCTCTTTCT GTTCCGGCAC
```

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32410      32420      32430      32440      32450      32460      32470      32480
TACCCGTCTC TTTCGCCCGC GCAGATTCGT ACTCAGCGAG CTGGGACATG CGGATAAGCT GCACAATGTG AGTATATTAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32490      32500      32510      32520      32530      32540      32550      32560
ACTAAGATAT ATAATCTATA TATATATAAA CATATATTCA TATATATAAT AATTATATGA TCTGTGTTTC CTGTACAATA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32570      32580      32590      32600      32610      32620      32630      32640
AACCATGAGT AACAAGACAA CTTAAATGAT AATGTTCGTT ATTCCGCTGC AGATGCCGGA TAACAATGCA CTGGCCGGAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32650      32660      32670      32680      32690      32700      32710      32720
TCTGCGATGG CATGGTGAAG GCGTGGGACA TCTATGCAAA GCCGCAGGCC GTCATTCTGT TCATCATCGA GGATGTGTCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32730      32740      32750      32760      32770      32780      32790      32800
TACAACATTT GCGACCAGCG GTTCCACGAG TTCTACATTC GCGAGACATA TCCGCACATC AAGGTGCTGC GCCGCACCCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32810      32820      32830      32840      32850      32860      32870      32880
CACCGAGGTC CATCGGGAGG GCAAACTGGG CCAGTCCAAG GAGTTGCTCC TGTAAGTGTA CATTAATACC TCTATTATCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32890      32900      32910      32920      32930      32940      32950      32960
TTCATCATTG AACTGTGTGC TCAACTCACA CGCTTGCCGC ACCACAGTGG ATCCCAAGAG GTGGCCGTCA TCTACTTCCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      32970      32980      32990      33000      33010      33020      33030      33040
AGCTGGCTAC GAGCCCGGAC ACTATCACTC GCAGGCGGAT TGGGATGCCC GCTATCTGAT GGAGACCTCG CTGGCCATCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33050      33060      33070      33080      33090      33100      33110      33120
AGTGCCCGTC GATTCACTAC CATTTGGCGG GCACCAAGAA GGTGCAGCAG GCATTGGCGC AGCCGGCGGT GCTCGAGCGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33130      33140      33150      33160      33170      33180      33190      33200
TTCATCAACG ATCCGGAGGA GATCAAGGCG GTGGGCAAGA TCTTCACGGG TCTCTACTCG CTGGACGACA ACGAGGCGGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33210      33220      33230      33240      33250      33260      33270      33280
CAATGCCAGC TACGAGATGG CGCTGCGCAC TCCGGAGAGA TTCGTGCTGA AGCCGCAGCG CGAGGGAGGC GGCAATAATG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33290      33300      33310      33320      33330      33340      33350      33360
TCTACGGCGT GGATATACCA GATGCACTAA AGGTGAGTTA TTCCACTTTG ATGATTATTT GAATTGTGAA CAGTTTTGTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33370      33380      33390      33400      33410      33420      33430      33440
ACAAATCCTT CTCCTTCTTT ATCCATCAGC GCATGTCACG CGTGGAGCGC TCCGCCTGGA TTCTGATGGA TCTGATTCAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33450      33460      33470      33480      33490      33500      33510      33520
CCGCCGCTGA CCAAGGGCTA TATGGTGCGG CCCGGTGGCG ATATGCCGCC CCAGATCGTG GACATGGTCT CCGAGCTGGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33530      33540      33550      33560      33570      33580      33590      33600
CATCTTCGGT GTGGTGATCG GAGATGCGGA GCACATTGTG CACAACTACC AGGCGGGACA CATGCTGCGC ACAAAGCTCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33610      33620      33630      33640      33650      33660      33670      33680
CAACCGCCAA CGAAGGAGGC GTGGCAGCCG GTCTGGGCGC CCTGGACAGT CCATATCTGA TCGATAGCGA TGACGAGGAT
End of distal DmGS copy
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33690      33700      33710      33720      33730      33740      33750      33760
GAGCAGAAGT AGGATCCTTT AACGAGATAA AGGCTTCGAA CGCCGGCATT CCACCTGTTG CCTCAGTCTT GTATAAATGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33770      33780      33790      33800      33810      33820      33830      33840
TTATTTTTGT AAAGCGTATT GTAAATTTTT TGCACTATTT TTTCCCCGTC CCAATGTCAA TAAATAAAAT GTTAAACTAG
```

Point of duplication

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33850      33860      33870      33880      33890      33900      33910      33920
AAATTCTGGT CAATATGCAA ACTATGCAAA ATATGCCAGG CTCAACGTTT TGATCAAAAA CATGAAAGTC CTTCGAAGTG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      33930      33940      33950      33960      33970      33980      33990      34000
AGTGTCTGAA TGCCCAGGCG GAGATCAAAA GCAAGGACAT ACAAATACAA TTGAGCGCGG CAAAGATCAA GCAAATGAGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34010      34020      34030      34040      34050      34060      34070      34080
AACGAACTGG TGACGCAGTG CAGCCGCCAG GACACGTGTC CCATCGACGG CAAAGGTGGC ATCTATAAGC TGAAGATCCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34090      34100      34110      34120      34130      34140      34150      34160
TGAATTGCCA GCATTCGAGG CGCCGTGCAG TTCCAACGGC TGGTTGACCA TACAGAAACG ATACGATGGC GCCGAGAACT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34170      34180      34190      34200      34210      34220      34230      34240
TCGATCGCGG TTGGAAGGAC TACAAGGATG GTTTCGGACG CGTGAGAGGT GAGTTTTTCA TCGGCCTGGA GAAAGTGCAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34250      34260      34270      34280      34290      34300      34310      34320
CTTATGACGC GGCAACGACG CCACGAACTG TACATTAAGC TGGGCAAGAT TGATGGCACC ACTAGCCACG CCCACTACGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34330      34340      34350      34360      34370      34380      34390      34400
CAACTTCGAA CTCGGCGGCG AGATTGAATC CTATGAGCTA AAGTCACTGG GCCGCTACAA TGGCACAGCC GGCGATTCTC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34410      34420      34430      34440      34450      34460      34470      34480
TGCGGCCACA TGAGCGCCAA AAGTTCACCA CCAACGACAA GGACAACGAT GCCTATCGGT TCAACTGTGC CGCCGACGAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34490      34500      34510      34520      34530      34540      34550      34560
TATGGTGGCT GGTGGTACTA CGATTGCGCC AAGAGGTAAG TACGCCCATT ATATACATAT GGTATATATA CATATGTACT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34570      34580      34590      34600      34610      34620      34630      34640
TGCATGTAAT ATTAATCGGC GGAGTAACTG CAGCATGCTC AATGGCAAGT TCTACAAGGA GGGCCGCTCG CGGAATGGAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34650      34660      34670      34680      34690      34700      34710      34720
AAACCAATGG CATCCTGTGG GGCTCCTGGC ACAACAACGA TTGGACCTAC TCGCTAACCT TCGTGGAGAT GATGATAAGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34730      34740      34750      34760      34770      34780      34790      34800
CCGCGAATAG ATTAGCTATT ACAGTAAGAA TAAAAAAATG GCACAAATAA ACGAAAAGGT CTCATTTGTA ATCATTACAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34810      34820      34830      34840      34850      34860      34870      34880
TCAATTTCAA TTAAATTAAA TAACGGGCAC TTACTAAACT TTTCTCGCCG TGTATGCCCC TGGTATTTTT TGATGCCATA
                                                                  Start of proximal DmGS copy
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34890      34900      34910      34920      34930      34940      34950      34960
TTCTGGCATA TTGCGGTATA TCGCCGGCTG GTATTTTTTG CCTAACCCGC TACGCACACA CTGCGAATAT TCGGCCAAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      34970      34980      34990      35000      35010      35020      35030      35040
ACATAGTTAT ATTTTTGTCT GTGGGCTTTT AGACAAATTG TAAACTGATA GCTCCCGACC CCAGCCGTTT AACGGTGTCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35050      35060      35070      35080      35090      35100      35110      35120
AAATTGACAG TGCAACCGGT GGAACAGGTG GAGCACAAGC AGCGGCACCC AGAAGGAAAC CAGCGATCCA GTCACAACAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35130      35140      35150      35160      35170      35180      35190      35200
TCGCTCGTTG AGAGCGCAGC GATAAGAGTT CCAGAGTTTC TAGAGTTTCC AGCGATACCC AGCAGTGGTG AGTAGCTAGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35210      35220      35230      35240      35250      35260      35270      35280
```

```
AGACAGTCGT CGCACTGCTC AGCAGTGCTC AGTGCCCAGT GAGCAGTTCC GGCGAGCAGA CCCCGTGAGA TACCGCGGCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35290      35300      35310      35320      35330      35340      35350      35360
AACAACTAAC AACTAAGATA CTCTGCCCCT TTTCTAGACG CAGCTGAGCC CCAGATCCCA GATCCCAGAT ACCAGCAGCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35370      35380      35390      35400      35410      35420      35430      35440
GACCGACCCA AGCAGAACGA TCATGTCCAG CGACGCCAAT ACGGCCGTGC TGCGCAACTG CATCCGCTTG CCGCTGGCGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35450      35460      35470      35480      35490      35500      35510      35520
AGGATGAGCT ACTGGAGGTG ACGGCCAAGG CCAAGGATTA TGCCATTATG CACGGCGCCG CCATGCGATC GAAGACGGCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35530      35540      35550      35560      35570      35580      35590      35600
TTCAGTCCGG ACTCGTTGAA TGTGGGTGAT ACAATGGGCA CTCGCTAGTG AGGGCTCGTA TTCATGCTAA GGTGTCTAGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35610      35620      35630      35640      35650      35660      35670      35680
TTCCGCTCAA GAAGCTCATT TCAATGGAAC AACATTCATT ACACATCCCT TATCGAGTTC AGTTCCTTTT TAAGTTATGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35690      35700      35710      35720      35730      35740      35750      35760
GTAATGCCAA ATGTGATTGG ATGCTTGAGT TTGTTTTTCA TTTTATTGAC GGTAGAGAAA GATTAGCAAT GTGACGGTAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35770      35780      35790      35800      35810      35820      35830      35840
AAGCTACACA TTGGCATTCA ATCGAGCTTC AAGCTAGCGT GCTGCTGCTT GAGCAACAGT TTGCCAGGAT ACTAAACATT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35850      35860      35870      35880      35890      35900      35910      35920
TTCATTTTCC AGTTCGCCCC CTTCGTGCTG GTGCCTTCCT CGTTTCCGCG CAAGGAGTTC GAAAAGGCGG TGGCCCTACA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      35930      35940      35950      35960      35970      35980      35990      36000
GCCGATCATC AACCGGCTGA TGCACAATGT GGCCCACGAC GAGGAGTTCA TCACGACGAC GCTGGCGGAG ACGATCAAAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36010      36020      36030      36040      36050      36060      36070      36080
TGGACGAGTT CACGGCCAAT CTGTTCAACA TCTATCGCAA GGTGCTGGCG CACGGATTCA CCCAGGTGAG ATTCCGTCAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36090      36100      36110      36120      36130      36140      36150      36160
CTGGCGTCCA TCCATCTCCG ATCCTCCGCC CTCTGCACAC CATCATCGTA TCACCCGCCA ACGTATCCAC CTCATTCCCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36170      36180      36190      36200      36210      36220      36230      36240
TCGGGAGGGA GCACATCTAC AGGCAGCCAG TGGGTTACTC GCTAGAATTT ATTGGGCCCC TCCACCCGCG GTCGCAATCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36250      36260      36270      36280      36290      36300      36310      36320
AATCGCCATT CCATAAATCG CCCGGCCATA GCCATAGCCG CAGTCAAAGC CCAGAAGTCG GCTGTAGTCG CTAATCTCGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36330      36340      36350      36360      36370      36380      36390      36400
GTGATTTATG CGAGCATCCT GGCCTCACAG CTCCTTATCG TTGGCTTGTA AATCGACGTA GGACAGCCCA TCCCTGACCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36410      36420      36430      36440      36450      36460      36470      36480
CCATAACGCT AGAACTTTGC CATTGCTCCT CCTACTTTGA TTTACATGCA TTTCCCCTAA CGAAACCAGC ACAGAGGCGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36490      36500      36510      36520      36530      36540      36550      36560
ATCCATTGAG CTCTTCCACG TCCACAAGAA GCTTTAATCT TTGTAGATCC TATCAGAAAG GTGACAGTTT TACAATTTGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36570      36580      36590      36600      36610      36620      36630      36640
ATAGTTATTG GCCCGGTAAC TGATAGGGTT TATTAAATAT AACAATGCCA GCGCTATTTT TACTTTATTA CTATTCAAAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
```

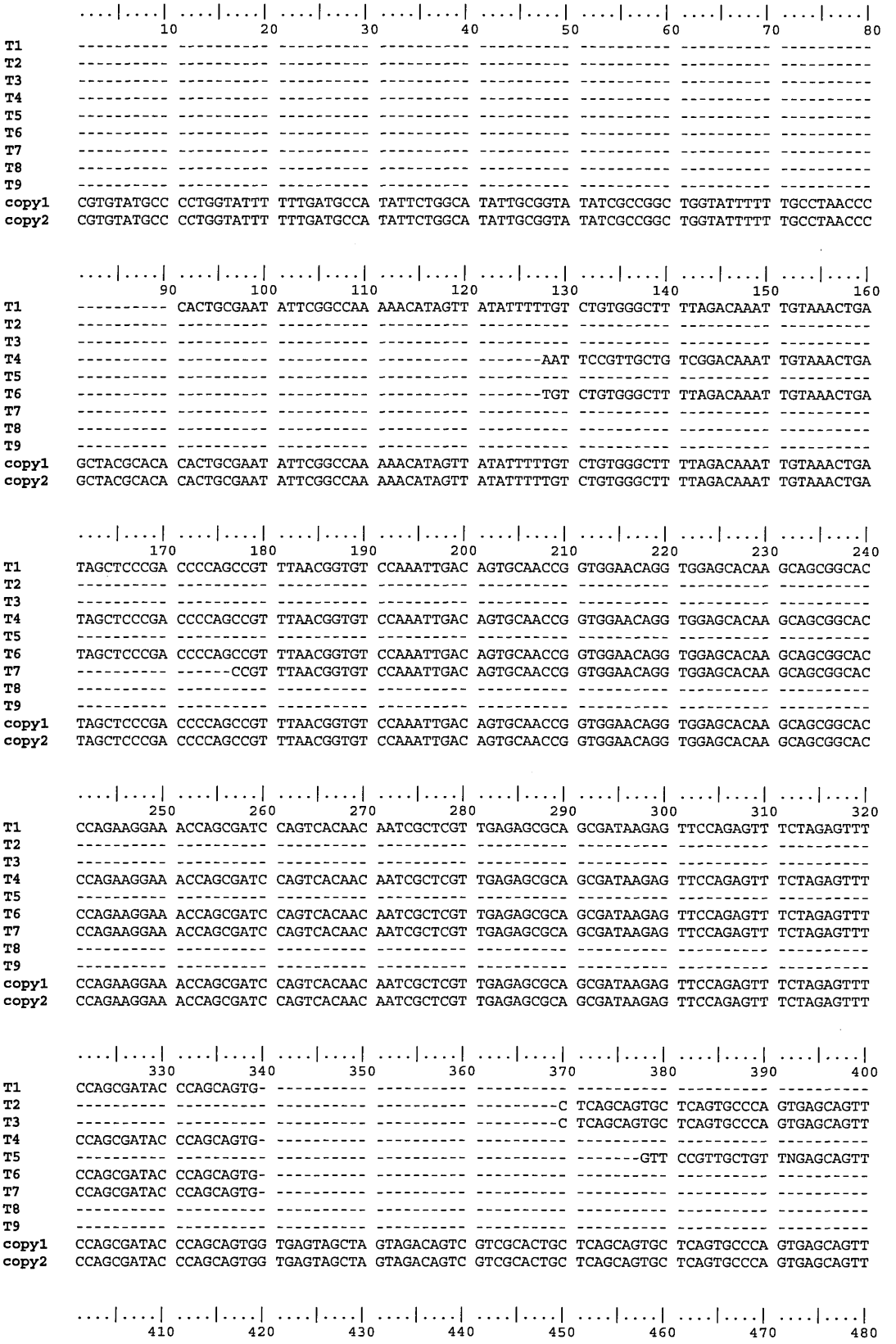
```
      36650      36660      36670      36680      36690      36700      36710      36720
TTTGTCTCGA ATCTTATCGA AGTCTCAATT GAAGCATTTA AATGCTAGAA TCAATTAAAC TATTGTTCAT TTTAAACTCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36730      36740      36750      36760      36770      36780      36790      36800
TTGATTTTTG CCAATGATTT TCAATGTTTT TACTTGCAAC TTAATCATTT TCATCACCCA ATATCAAGTT TAGTTGCGAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36810      36820      36830      36840      36850      36860      36870      36880
TTTTGCTCTC TGTGCAAAAA TTTAGTTCAC GATCTGTGCT TTTCTCGTCA CCACTCCTTC AAAAATTTTA AAACCTAACC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36890      36900      36910      36920      36930      36940      36950      36960
ACGCAATTGT ACAAACTCTC GATCTCAAAC GCAATCTGTT GCTCCATTAT CCGTCTCTCA TTGCATCGCT GTGCTACGAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      36970      36980      36990      37000      37010      37020      37030      37040
AAATGATTTT GAATCAAAGA AAACATCGCT GGGGATGCTG CGCAGCGATC TGATGCTGGA GTCCGGCTGT CCCGAGCTGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37050      37060      37070      37080      37090      37100      37110      37120
CGCCACGGGC GCTGCGAACG GCGGCCGGTG AGGATCGTGG TCAGGATGCG GGCGCTGCTG TGGGGCAAAT AGCTGGAGCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37130      37140      37150      37160      37170      37180      37190      37200
AACGGAGCTG CAGGAGTCGG AACTGCCGCT GGCACGGGCA GCAAGGAGGA GGAGCACAGG GAGGTCCAGC TAAGCCGGGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37210      37220      37230      37240      37250      37260      37270      37280
CACCAAAGAG CCGGAACGGA GGGCGACCGG CGCCCAGTCC GCTTACTGCT GCTGGAAGCA AGTCGAGATC AATACGATAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37290      37300      37310      37320      37330      37340      37350      37360
CATCGGGCTT TGGCCACTTG GGACCAGCAA GTAAAACCAT ACAAAGGTGC TGTTCACGCG CACAGTGCAT CCATCAGTGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37370      37380      37390      37400      37410      37420      37430      37440
AACTCCGCAC CACGCCATCA TCAGCAACAC ACCCCCTAAA AAAAAAAACC CAGTATTCCT TCCATGACAC GTCACTCATC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37450      37460      37470      37480      37490      37500      37510      37520
TCATCTCATC GCATTTTTGC ATTCCTGCAT CGCTTGCATC CTCCGCCTAA CCGCACAAAA TCACACACTA ACTATTTTCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37530      37540      37550      37560      37570      37580      37590      37600
ATACCCGAAT CAAATATTAA CTATATTTGA ATCCTTAACT TTGTCAATCG AAAGAATCTT TTGCATCTGA AAAGAAAGAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37610      37620      37630      37640      37650      37660      37670      37680
CACTAAAGTC AAAGATAGAT CGGAATTAAA ACGAAGTTGC AACAAGAAGT GCATAATCAG ATATTTCGAC TGAAATAAGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37690      37700      37710      37720      37730      37740      37750      37760
TATTTAAAAT ATTTAAAAAA TATTTCAAAC ATATCTTGTA GTGTGTGCTT TTGTGTGTGT TGCTTTGTAA AATCCTTTGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37770      37780      37790      37800      37810      37820      37830      37840
ACTTTAGTTG AAGTTGATTT TGTTCCGCTG TTCCTTAGAA ATGTATTTAT GATTAACCAT AGCTACAAAC GTTGCTCCTC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37850      37860      37870      37880      37890      37900      37910      37920
CAGATTGTTG CCTAGAGTTT TCATTTGCTG ACGCTGCTGT CGCCTGTCGC CCGTTGCCTC CTGCCTGCCT GCTCTGCTCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      37930      37940      37950      37960      37970      37980      37990      38000
TGCCTCCTGG CATCCCGCAT CCTGCATCCT GGCACCTGCC AACCGACACC TGCTCCTCCC GACCCAACCG TAGCATCTAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38010      38020      38030      38040      38050      38060      38070      38080
CCCGCCCGCT AATAATCTTC ACTTCCGTCA GTGATTGTCC GTTCTGCTGT CGCTTACTCT ACCAACTCTA TAGTAGGTCA
```

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38090      38100      38110      38120      38130      38140      38150      38160
GGGGTTAGCC AATATATATG CATATATATC TTTCTAGAAG ATTTGAAGAT AGCGATTTCA CAAGATCTGA TAAAGTGTTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38170      38180      38190      38200      38210      38220      38230      38240
GGATGATACA TCACATCAAT TTGTCATTTG TAATCAGCAA AAGGCGCTAA TCAAAAATTT CACCACTTGT ACGCTTGGGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38250      38260      38270      38280      38290      38300      38310      38320
TATTAAAACC CTAGAAGTTT CTGGAGGTTG AATTAAACAA TTTAATTGTT TGAAAGAACA AATGTAGTTT TTACTTGCAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38330      38340      38350      38360      38370      38380      38390      38400
AGATATTAAA CTCACTCAAC ATTTTGGTCA TACGTCGTTT TAGACGAATC ACACCTGTAG TTAGGAAGCC CCACATGGAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38410      38420      38430      38440      38450      38460      38470      38480
AGCCCACAGC TGTACTGCAG ATAGGGTTAC AGATAGGGAT ACGCTGATAT AGATACGCCT GGTAATTGTT TATTTAAAGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38490      38500      38510      38520      38530      38540      38550      38560
CGAGGCGTGT GCGATAACAG ACCCAATAAG TGCATTTGTG GGTGCAAAAT GTTTGTCTTC AATTACAAAT TCAAGACTCG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38570      38580      38590      38600      38610      38620      38630      38640
AATTTATGAA ATGCCAGCAG AACTCAAAAC ACTTTCATGC ACGTTTCGAG TTTTTAGCTG TTTGTCTATT TGTTTAATTG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38650      38660      38670      38680      38690      38700      38710      38720
CCGGCAATGG TGACGTAAAC ATAATTATAA ACAGATTATA TCAGAGCATG TTTATAAAAT AGTCAAATTA CCAGCTGCAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38730      38740      38750      38760      38770      38780      38790      38800
ATGGGCTAAT AACTTGTTAT TACTTCGTCA TGCTTAGATA CATATGGAAA ATAGGTATTA GTAGTAGTAT TATTATTATT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38810      38820      38830      38840      38850      38860      38870      38880
ATTTCATTCA ATAACAGCTT GTAGTCATAT CGGAAATATT TAAACAATTT GAAACTGGCG GCGTTTGTGG CGAAATATTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38890      38900      38910      38920      38930      38940      38950      38960
GGAGTCTTTC ATGGCCCTAA AAATTATAGA GTTAATAGCA TTGGGAGATG CATCGTTTCT TGCACCATTT CTCGCCTCTC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      38970      38980      38990      39000      39010      39020      39030      39040
TCTCTCTCTC CGTCTCTTTT TGAATATATA TAATTAAAAA TTCGCACCTG CCTCATCCTA CATATATATA CATATATATA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39050      39060      39070      39080      39090      39100      39110      39120
TATATATATA TATGCACGTA TATATCTATG GACGAGGATC GTAAACCATA CCTAGACATA TTGAATCGCT TTACCGGCTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39130      39140      39150      39160      39170      39180      39190      39200
AGCAGCTAAT CTTTAAGGGA CGCACCTAAT GTATATATAT CTTGTATAAC GATCATAATT GCAGGCATAT AGCATCGGTC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39210      39220      39230      39240      39250      39260      39270      39280
TGTTGCGATC CGACTATATG GCACACGTGT CCGAGGGCTG TGCGATAAAG CAGGTCGAGA TCAACACGGT GGCCTCCAGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39290      39300      39310      39320      39330      39340      39350      39360
TTCGCTGGCA TAGCCACCCA ATTGGTGCCG CAGCAGAAGT ATTAAGCGCT CTTCTCTCTA TCTCTTTCTA CATATAATAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39370      39380      39390      39400      39410      39420      39430      39440
ATATAAATAT ATATCTATCA AACTACATAT ATTCCTTTTA CCAAACGCAC GAATGCACAA TGCACACACC TAAAAGCAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39450      39460      39470      39480      39490      39500      39510      39520
AGCACACTAC ACACCTGTCC TAATATCCTT CCACCCAACT CCTTCTCCTT CTCACTGATC ACTTATCCAC CTACAGAGGT
```

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39530      39540      39550      39560      39570      39580      39590      39600
GCCATCTCTA TCGTTTGTTG TCTCTTTCTG CATGTGTTTT TTTTTTTTTT AATTGGTTAT GGCTCTATCC GTCTCTTTCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39610      39620      39630      39640      39650      39660      39670      39680
ACACGCTGTA GCTTCATAAC CCCACTTCCA CAATTTCCCT AACCATAAAC ATACATATGT ATATCCATCC GCCCATTTGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39690      39700      39710      39720      39730      39740      39750      39760
GAGGTTTTGG TCTCAGTAAT TAACCCAAAT TTCAACCCCC CCCCCCCCAT TACCACCGAA CATTTGTTGT TGTTTGTTGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39770      39780      39790      39800      39810      39820      39830      39840
TGTTGTTGGT CGCTGAAAGA CTGCGGAATT ATGCTAATGT CTGTCTCTTT CTGTTCCGGC ACTACCCGTC TCTTTCGCCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39850      39860      39870      39880      39890      39900      39910      39920
GCGCAGATTC GTACTCAGCG AGCTGGGACA TGCGGATAAG CTGCACAATG TGAGTATATT AAACTAAGAT ATATAATCTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      39930      39940      39950      39960      39970      39980      39990      40000
TATATATATA AACATATATT CATATATATA ATAATTATAT GATCTGTGTT TCCTGTACAA TAAACCATGA GTAACAAGAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40010      40020      40030      40040      40050      40060      40070      40080
AACTTAAATG ATAATGTTCG TTATTCCGCT GCAGATGCCG GATAACAATG CACTGGCCGG ACTCTGCGAT GGCATGGTGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40090      40100      40110      40120      40130      40140      40150      40160
AGGCGTGGGA CATCTATGCA AAGCCGCAGG CCGTCATTCT GTTCATCATC GAGGATGTGT CGTACAACAT TTGCGACCAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40170      40180      40190      40200      40210      40220      40230      40240
CGGTTCCACG AGTTCTACAT TCGCGAGACA TATCCGCACA TCAAGGTGCT GCGCCGCACC CTCACCGAGG TCCATCGGGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40250      40260      40270      40280      40290      40300      40310      40320
GGGCAAACTG GGCCAGTCCA AGGAGTTGCT CCTGTAAGTG TACATTAATA CCTCTATTAT CATTCATCAT TGAACTGTGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40330      40340      40350      40360      40370      40380      40390      40400
GCTCAACTCA CACGCTTGCC GCACCACAGT GGATCCCAAG AGGTGGCCGT CATCTACTTC CGAGCTGGCT ACGAGCCCGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40410      40420      40430      40440      40450      40460      40470      40480
ACACTATCAC TCGCAGGCGG ATTGGGATGC CCGCTATCTG ATGGAGACCT CGCTGGCCAT CAAGTGCCCG TCGATTCACT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40490      40500      40510      40520      40530      40540      40550      40560
ACCATTTGGC GGGCACCAAG AAGGTGCAGC AGGCATTGGC GCAGCCGGCG GTGCTCGAGC GTTTCATCAA CGATCCGGAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40570      40580      40590      40600      40610      40620      40630      40640
GAGATCAAGG CGGTGGGCAA GATCTTCACG GGTCTCTACT CGCTGGACGA CAACGAGGCG GGCAATGCCA GCTACGAGAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40650      40660      40670      40680      40690      40700      40710      40720
GGCGCTGCGC ACTCCGGAGA GATTCGTGCT GAAGCCGCAG CGCGAGGGAG GCGGCAATAA TGTCTACGGC GTGGATATAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40730      40740      40750      40760      40770      40780      40790      40800
CAGATGCACT AAAGGTGAGT TATTCCACTT TGATGATTAT TTGAATTGTG AACAGTTTTG TAACAAATCC TTCTCCTTCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40810      40820      40830      40840      40850      40860      40870      40880
TTATCCATCA GCGCATGTCA CGCGTGGAGC GCTCCGCCTG GATTCTGATG GATCTGATTC ATCCGCCGCT GACCAAGGGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40890      40900      40910      40920      40930      40940      40950      40960
```

```
TATATGGTGC GGCCCGGTGG CGATATGCCG CCCCAGATCG TGGACATGGT CTCCGAGCTG GGCATCTTCG GTGTGGTGAT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      40970      40980      40990      41000      41010      41020      41030      41040
CGGAGATGCG GAGCACATTG TGCACAACTA CCAGGCGGGA CACATGCTGC GCACAAAGCT CTCAACCGCC AACGAAGGAG
                                                                  End of proximal DmGS copy
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41050      41060      41070      41080      41090      41100      41110      41120
GCGTGGCAGC CGGTCTGGGC GCCCTGGACA GTCCATATCT GATCGATAGC GATGACGAGG ATGAGCAGAA GTAGGATCCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41130      41140      41150      41160      41170      41180      41190      41200
TTAACGAGAT AAAGGCTTCG AACGCCGGCA TTCCACCTGT TGCCTCAGTC TTGTATAAAT GATTATTTTT GTAAAGCGTA
                                                                  End of duplicated region
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41210      41220      41230      41240      41250      41260      41270      41280
TTGTAAATTT TTTGCACTAT TTTTTCCCCG TCCCAATGTC AATAAATAAA ATGTTAAACT AGAAATTCTG GTCAATATGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41290      41300      41310      41320      41330      41340      41350      41360
AAACTGGATT TGGGGCTAAA GCTTTGCTAC AGTTCATGTA AACTCTTAGT ATCTAAGAAC TATCACCGTG GAGTAGAATG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41370      41380      41390      41400      41410      41420      41430      41440
GGAGTAGACT TGGAGTTGGG CCAATGCTGT GCACATCCCA GCTAGCGTTC ACTACCGCAC CGTCTTGCGG TAGGTGGCGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41450      41460      41470      41480      41490      41500      41510      41520
CCATCAGGCG GTCAACCATG CTGTCAATGG GATGGATGAT AAGCGGTATG CTGGCCAGAC CCACCAGGGT GGGCACCGTC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41530      41540      41550      41560      41570      41580      41590      41600
TTTAGCACGG GCATAGGCGC TTTGCTCAGC AGTGTTTTGG TGGCCCAGGT GATCCTAAAA CGATATAAAT GAGCATTGCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41610      41620      41630      41640      41650      41660      41670      41680
TATTATGTGA ACATAAACTC ATAAAAATGG GTATGATCTT GGAGAGAACG ATGACTTCAT CGTGTATCTA CAAATATACA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41690      41700      41710      41720      41730      41740      41750      41760
ATAAACCTGT TTTTGCAACA TTGCACATGC CATTCTACTA CGTATATGCA CTCTATAGTA TAATTTTTAC TATGCTTATA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41770      41780      41790      41800      41810      41820      41830      41840
TAAGATAAGT ATTATGCAAC TAATGGAAAA AGGGAATCTT TAAGACGCCT CATTCCCTAC TTCCAACTAT GTAAATACAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41850      41860      41870      41880      41890      41900      41910      41920
CCATTAAGTC TACGTTTTTG GGCAATTTTC TGGCAAGGGT TGAACGCACC GATTGATGAC CAGGCCGGGA ATGGCGACGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      41930      41940      41950      41960      41970      41980      41990      42000
AGGCCAGCAT CTGCCAGCAG AAGACGTCGC CGCCCTTGAC GGCCACTTCT CTGGAGGAAG CGCCCTCCAT TTGGAGGCGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      42010      42020      42030      42040      42050      42060      42070      42080
AGTGCCTTGT CAAAAGTGTC CGTGCAGACG TATCCGATGG CCATGCCGTA GGATGCGGCC ACCAGAGACT TGGGCACCAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      42090      42100      42110      42120      42130      42140      42150      42160
CGGGCGGAAG GACTCGCCAA TCTCGTTGCT GTAGCCTGCA AAGCCGTAAC AATTAAACTC ATATATTTCG ATGGATCTGG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      42170      42180      42190      42200      42210      42220      42230      42240
GCAGATTCTG CTTGGGAGGA ACACTCACCC ATATAGCGAA TGAAGGTGTC CCGGTAGATG TCCACCTCCT TCAAAGACGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      42250      42260      42270      42280      42290      42300      42310      42320
GGAAGTTTTC TTGGTGCCCT TATCCCGGCG GATGAACTGT TTGTCCTCCG ACATTGTCTA TTTTTTTTAT TTTTATTTTT
```

9.2 Appendix B: Sequence alignment of all nine DmGS cDNAs against the genomic sequence



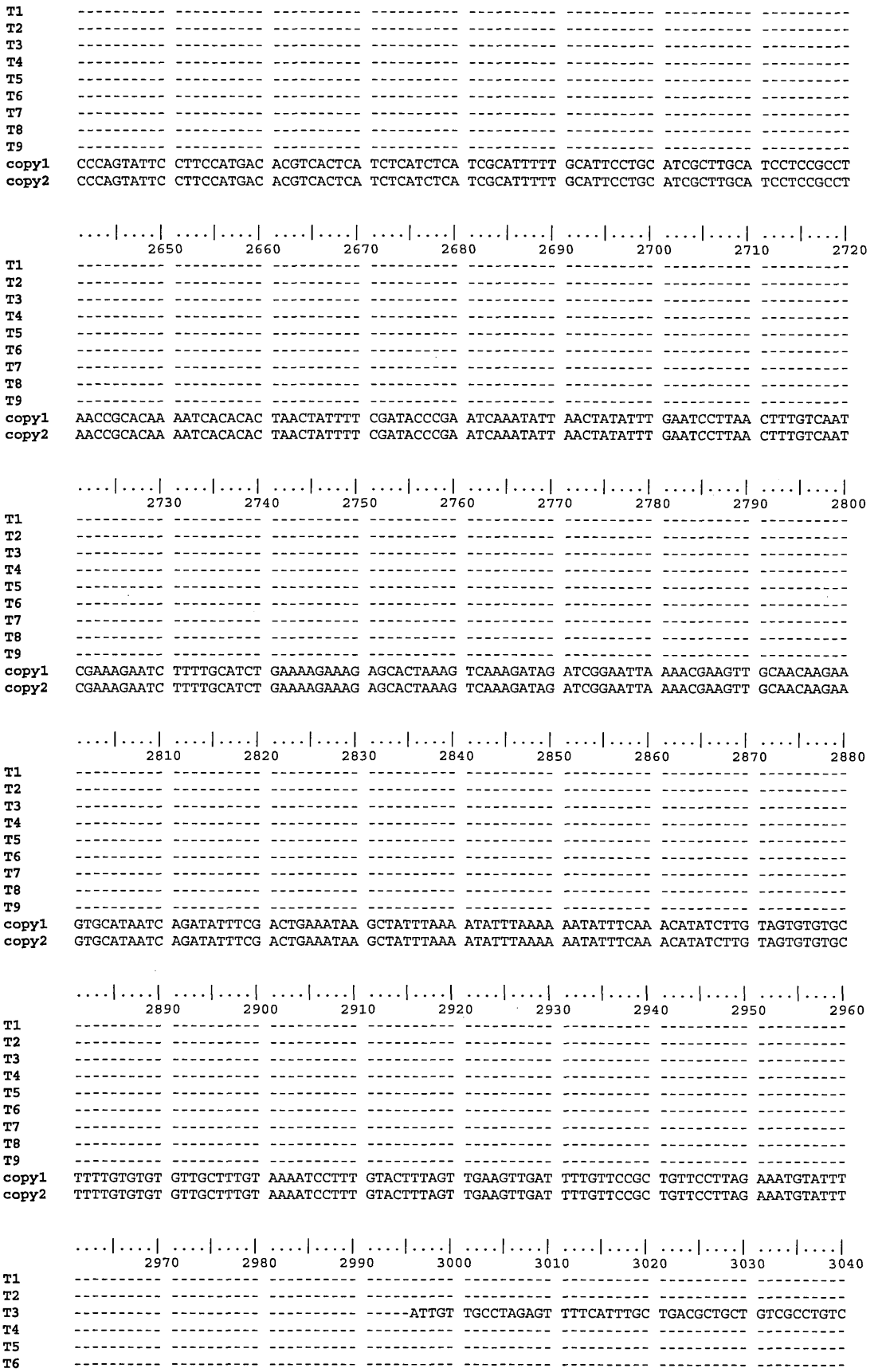
```
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------A CGCAGCTGAG
T2      CCGGCGAGCA GACCCCGTGA GATACCGCGG CGAACAACTA ACAACTAAGA TACTCTGCCC CTTTTCTAGA CGCAGCTGAG
T3      CCGGCGAGCA GACCCCGTGA GATACCGCGG CGAACAACTA ACAACTAAGA TACTCTGCCC CTTTTCTAGA CGCAGCTGAG
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------A CGCAGCTGAG
T5      CCGGCgagCA GACCCCGTGA GATACCGCGG CGAACAACTA ACAACTAAGA TACTCTGCCC CTTTTCTAGA CGCAGCTGAG
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------A CGCAGCTGAG
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------A CGCAGCTGAG
T8      CCGGCGAGCA GACCCCGTGA GATACCGCGG CGAACAACTA ACAACTAAGA TACTCTGCCC CTTTTCTAGA CGCAGCTGAG
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CCGGCGAGCA GACCCCGTGA GATACCGCGG CGAACAACTA ACAACTAAGA TACTCTGCCC CTTTTCTAGA CGCAGCTGAG
copy2   CCGGCGAGCA GACCCCGTGA GATACCGCGG CGAACAACTA ACAACTAAGA TACTCTGCCC CTTTTCTAGA CGCAGCTGAG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                      490        500        510        520        530        540        550        560
T1      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T2      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T3      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T4      CCCCAGATCC CAGATCCCAG ATACCAGCAG TAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T5      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T6      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T7      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T8      CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT
copy2   CCCCAGATCC CAGATCCCAG ATACCAGCAG CAGACCGACC CAAGCAGAAC GATCATGTCC AGCGACGCCA ATACGGCCGT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                      570        580        590        600        610        620        630        640
T1      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T2      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T3      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T4      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T5      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T6      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T7      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T8      GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA
copy2   GCTGCGCAAC TGCATCCGCT TGCCGCTGGC GGAGGATGAG CTACTGGAGG TGACGGCCAA GGCCAAGGAT TATGCCATTA

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                      650        660        670        680        690        700        710        720
T1      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T2      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T3      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T4      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T5      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T6      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T7      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T8      TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AAT------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AATGTGGGTG ATACAATGGG CACTCGCTAG
copy2   TGCACGGCGC CGCCATGCGA TCGAAGACGG CCTTCAGTCC GGACTCGTTG AATGTGGGTG ATACAATGGG CACTCGCTAG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                      730        740        750        760        770        780        790        800
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGAGGGCTCG TATTCATGCT AAGGTGTCTA GTTTCCGCTC AAGAAGCTCA TTTCAATGGA ACAACATTCA TTACACATCC
copy2   TGAGGGCTCG TATTCATGCT AAGGTGTCTA GTTTCCGCTC AAGAAGCTCA TTTCAATGGA ACAACATTCA TTACACATCC

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                      810        820        830        840        850        860        870        880
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CTTATCGAGT TCAGTTCCTT TTTAAGTTAT GTGTAATGCC AAATGTGATT GGATGCTTGA GTTTGTTTTT CATTTTATTG
copy2   CTTATCGAGT TCAGTTCCTT TTTAAGTTAT GTGTAATGCC AAATGTGATT GGATGCTTGA GTTTGTTTTT CATTTTATTG

                 890        900        910        920        930        940        950        960
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   ACGGTAGAGA AAGATTAGCA ATGTGACGGT ATAAGCTACA CATTGGCATT CAATCGAGCT TCAAGCTAGC GTGCTGCTGC
copy2   ACGGTAGAGA AAGATTAGCA ATGTGACGGT ATAAGCTACA CATTGGCATT CAATCGAGCT TCAAGCTAGC GTGCTGCTGC

                 970        980        990       1000       1010       1020       1030       1040
T1      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T2      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T3      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T4      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T5      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T6      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T7      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T8      ---------- ---------- ---------- ---------- ----TTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TTGAGCAACA GTTTGCCAGG ATACTAAACA TTTTCATTTT CCAGTTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG
copy2   TTGAGCAACA GTTTGCCAGG ATACTAAACA TTTTCATTTT CCAGTTCGCC CCCTTCGTGC TGGTGCCTTC CTCGTTTCCG

                1050       1060       1070       1080       1090       1100       1110       1120
T1      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T2      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T3      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T4      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T5      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T6      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T7      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T8      CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT
copy2   CGCAAGGAGT TCGAAAAGGC GGTGGCCCTA CAGCCGATCA TCAACCGGCT GATGCACAAT GTGGCCCACG ACGAGGAGTT

                1130       1140       1150       1160       1170       1180       1190       1200
T1      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T2      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T3      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T4      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T5      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGATGCTGG
T6      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T7      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T8      CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG
copy2   CATCACGACG ACGCTGGCGG AGACGATCAA AGTGGACGAG TTCACGGCCA ATCTGTTCAA CATCTATCGC AAGGTGCTGG

                1210       1220       1230       1240       1250       1260       1270       1280
T1      CGCACGGATT CACCCAGG-- ---------- ---------- ---------- ---------- ---------- ----------
T2      CGCACGGATT CACCCAGG-- ---------- ---------- ---------- ---------- ---------- ----------
T3      CGCACGGATT CACCCAGg-- ---------- ---------- ---------- ---------- ---------- ----------
T4      CACACGGATT CACCCAGG-- ---------- ---------- ---------- ---------- ---------- ----------
T5      CGCACGGATT CACCCAG--- ---------- ---------- ---------- ---------- ---------- ----------
T6      CGCACGGATT CACCCAGg-- ---------- ---------- ---------- ---------- ---------- ----------
T7      CGCACGGATT CACCCAGg-- ---------- ---------- ---------- ---------- ---------- ----------
T8      CGCACGGATT CACCCAGg-- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CGCACGGATT CACCCAGGTG AGATTCCGTC AGCTGGCGTC CATCCATCTC CGATCCTCCG CCCTCTGCAC ACCATCATCG
copy2   CGCACGGATT CACCCAGGTG AGATTCCGTC AGCTGGCGTC CATCCATCTC CGATCCTCCG CCCTCTGCAC ACCATCATCG
```

```
          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  1290       1300       1310       1320       1330       1340       1350       1360
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     TATCACCCGC CAACGTATCC ACCTCATTCC CATCGGGAGG GAGCACATCT ACAGGCAGCC AGTGGGTTAC TCGCTAGAAT
copy2     TATCACCCGC CAACGTATCC ACCTCATTCC CATCGGGAGG GAGCACATCT ACAGGCAGCC AGTGGGTTAC TCGCTAGAAT

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  1370       1380       1390       1400       1410       1420       1430       1440
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     TTATTGGGCC CCTCCACCCG CGGTCGCAAT CTAATCGCCA TTCCATAAAT CGCCCGGCCA TAGCCATAGC CGCAGTCAAA
copy2     TTATTGGGCC CCTCCACCCG CGGTCGCAAT CTAATCGCCA TTCCATAAAT CGCCCGGCCA TAGCCATAGC CGCAGTCAAA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  1450       1460       1470       1480       1490       1500       1510       1520
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     GCCCAGAAGT CGGCTGTAGT CGCTAATCTC GCGTGATTTA TGCGAGCATC CTGGCCTCAC AGCTCCTTAT CGTTGGCTTG
copy2     GCCCAGAAGT CGGCTGTAGT CGCTAATCTC GCGTGATTTA TGCGAGCATC CTGGCCTCAC AGCTCCTTAT CGTTGGCTTG

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  1530       1540       1550       1560       1570       1580       1590       1600
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     TAAATCGACG TAGGACAGCC CATCCCTGAC CCCCATAACG CTAGAACTTT GCCATTGCTC CTCCTACTTT GATTTACATG
copy2     TAAATCGACG TAGGACAGCC CATCCCTGAC CCCCATAACG CTAGAACTTT GCCATTGCTC CTCCTACTTT GATTTACATG

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  1610       1620       1630       1640       1650       1660       1670       1680
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     CATTTCCCCT AACGAAACCA GCACAGAGGC GAATCCATTG AGCTCTTCCA CGTCCACAAG AAGCTTTAAT CTTTGTAGAT
copy2     CATTTCCCCT AACGAAACCA GCACAGAGGC GAATCCATTG AGCTCTTCCA CGTCCACAAG AAGCTTTAAT CTTTGTAGAT

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  1690       1700       1710       1720       1730       1740       1750       1760
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CCTATCAGAA AGGTGACAGT TTTACAATTT GAATAGTTAT TGGCCCGGTA ACTGATAGGG TTTATTAAAT ATAACAATGC
copy2   CCTATCAGAA AGGTGACAGT TTTACAATTT GAATAGTTAT TGGCCCGGTA ACTGATAGGG TTTATTAAAT ATAACAATGC

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    1770       1780       1790       1800       1810       1820       1830       1840
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CAGCGCTATT TTTACTTTAT TACTATTCAA ACTTTGTCTC GAATCTTATC GAAGTCTCAA TTGAAGCATT TAAATGCTAG
copy2   CAGCGCTATT TTTACTTTAT TACTATTCAA ACTTTGTCTC GAATCTTATC GAAGTCTCAA TTGAAGCATT TAAATGCTAG

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    1850       1860       1870       1880       1890       1900       1910       1920
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AATCAATTAA ACTATTGTTC ATTTTAAACT CGTTGATTTT TGCCAATGAT TTTCAATGTT TTTACTTGCA ACTTAATCAT
copy2   AATCAATTAA ACTATTGTTC ATTTTAAACT CGTTGATTTT TGCCAATGAT TTTCAATGTT TTTACTTGCA ACTTAATCAT

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    1930       1940       1950       1960       1970       1980       1990       2000
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TTTCATCACC CAATATCAAG TTTAGTTGCG AATTTTGCTC TCTGTGCAAA AATTTAGTTC ACGATCTGTG CTTTTCTCGT
copy2   TTTCATCACC CAATATCAAG TTTAGTTGCG AATTTTGCTC TCTGTGCAAA AATTTAGTTC ACGATCTGTG CTTTTCTCGT

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    2010       2020       2030       2040       2050       2060       2070       2080
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CACCACTCCT TCAAAAATTT TAAAACCTAA CCACGCAATT GTACAAACTC TCGATCTCAA ACGCAATCTG TTGCTCCATT
copy2   CACCACTCCT TCAAAAATTT TAAAACCTAA CCACGCAATT GTACAAACTC TCGATCTCAA ACGCAATCTG TTGCTCCATT

            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    2090       2100       2110       2120       2130       2140       2150       2160
T1      ---------- ---------- ---------- ---------- ---------- -AAAACATCG CTGGGGATGC TGCGCAGCGA
T2      ---------- ---------- ---------- ---------- ---------- -AAAACATCG CTGGGGATGC TGCGCAGCGA
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- -AAAACATCG CTGGGGATGC TGCGCAGCGA
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- -AAAACATCG CTGGGGATGC TGCGCAGCGA
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
copy1   ATCCGTCTCT CATTGCATCG CTGTGCTACG AAAAATGATT TTGAATCAAA GAAAACATCG CTGGGGATGC TGCGCAGCGA
copy2   ATCCGTCTCT CATTGCATCG CTGTGCTACG AAAAATGATT TTGAATCAAA GAAAACATCG CTGGGGATGC TGCGCAGCGA

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                2170       2180       2190       2200       2210       2220       2230       2240
T1      TCTGATGCTG GAGTCCGGCT GTCCCGAGCT GTCGCCACGG GCGCTGCGAA CGGCGGCCGG TGAGGATCGT GGTCAGGATG
T2      TCTGATGCTG GAGTCCGGCT GTCCCGAGCT GTCGCCACGG GCGCTGCGAA CGGCGGCCGG TGAGGATCGT GGTCAGGATG
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      TCTGATGCTG GAGTCCGGCT GTCCCGAGCT GTCGCCACGG GCGCTGCGAA CGGCGGCCGG TGAGGATCGT GGTCAGGATG
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      TCTGATGCTG GAGTCCGGCT GTCCCGAGCT GTCGCCACGG GCGCTGCGAA CGGCGGCCGG TGAGGATCGT GGTCAGGATG
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TCTGATGCTG GAGTCCGGCT GTCCCGAGCT GTCGCCACGG GCGCTGCGAA CGGCGGCCGG TGAGGATCGT GGTCAGGATG
copy2   TCTGATGCTG GAGTCCGGCT GTCCCGAGCT GTCGCCACGG GCGCTGCGAA CGGCGGCCGG TGAGGATCGT GGTCAGGATG

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                2250       2260       2270       2280       2290       2300       2310       2320
T1      CGGGCGCTGC TGTGGGGCAA ATAGCTGGAG CTAACGGAGC TGCAGGAGTC GGAACTGCCG CTGGCACGGG CAGCAAGGAG
T2      CGGGCGCTGC TGTGGGGCAA ATAGCTGGAG CTAACGGAGC TGCAGGAGTC GGAACTGCCG CTGGCACGGG CAGCAAGGAG
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      CGGGCGCTGC TGTGGGGCAA ATAGCTGGAG CTAACGGAGC TGCAGGAGTC GGAACTGCCG CTGGCACGGG CAGCAAGGAG
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      CGGGCGCTGC TGTGGGGCAA ATAGCTGGAG CTAACGGAGC TGCAGGAGTC GGAACTGCCG CTGGCACGGG CAGCAAGGAG
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CGGGCGCTGC TGTGGGGCAA ATAGCTGGAG CTAACGGAGC TGCAGGAGTC GGAACTGCCG CTGGCACGGG CAGCAAGGAG
copy2   CGGGCGCTGC TGTGGGGCAA ATAGCTGGAG CTAACGGAGC TGCAGGAGTC GGAACTGCCG CTGGCACGGG CAGCAAGGAG

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                2330       2340       2350       2360       2370       2380       2390       2400
T1      GAGGAGCACA GGGAGGTCCA GCTAAGCCGG GTCACCAAAG AGCCGGAACG GAGGGCGACC GGCGCCCAGT CCGCTTACTG
T2      GAGGAGCACA GGGAGGTCCA GCTAAGCCGG GTCACCAAAG AGCCGGAACG GAGGGCGACC GGCGCCCAGT CCGCTTACTG
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      GAGGAGCACA GGGAGGTCCA GCTAAGCCGG GTCACCAAAG AGCCGGAACG GAGGGCGACC GGCGCCCAGT CCGCTTACTG
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      GAGGAGCACA GGGAGGTCCA GCTAAGCCGG GTCACCAAAG AGCCGGAACG GAGGGCGACC GGCGCCCAGT CCGCTTACTG
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GAGGAGCACA GGGAGGTCCA GCTAAGCCGG GTCACCAAAG AGCCGGAACG GAGGGCGACC GGCGCCCAGT CCGCTTACTG
copy2   GAGGAGCACA GGGAGGTCCA GCTAAGCCGG GTCACCAAAG AGCCGGAACG GAGGGCGACC GGCGCCCAGT CCGCTTACTG

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                2410       2420       2430       2440       2450       2460       2470       2480
T1      CTGCTGGAAG CAAGTCGAGA TCAATACGAT AGCATCGGGC TTTGGCCACT TGGGACCAGC AAGTAAAACC ATACAAAG--
T2      CTGCTGGAAG CAAGTCGAGA TCAATACGAT AGCATCGGGC TTTGGCCACT TGGGACCAGC AAGTAAAACC ATACAAAG--
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      CTGCTGGAAG CAAGTCGAGA TCAATACGAT AGCATCGGGC TTTGGCCACT TGGGACCAGC AAGTAAAACC ATACAAAG--
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      CTGCTGGAAG CAAGTCGAGA TCAATACGAT AGCATCGGGC TTTGGCCACT TGGGACCAGC AA-------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- --------AT AGCATCGGGC TTTGGCCACT TGGGACCAGC AAGTAAAACC ATACAAAG--
copy1   CTGCTGGAAG CAAGTCGAGA TCAATACGAT AGCATCGGGC TTTGGCCACT TGGGACCAGC AAGTAAAACC ATACAAAGGT
copy2   CTGCTGGAAG CAAGTCGAGA TCAATACGAT AGCATCGGGC TTTGGCCACT TGGGACCAGC AAGTAAAACC ATACAAAGGT

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                2490       2500       2510       2520       2530       2540       2550       2560
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GCTGTTCACG CGCACAGTGC ATCCATCAGT GCAACTCCGC ACCACGCCAT CATCAGCAAC ACACCCCCTA AAAAAAAAAA
copy2   GCTGTTCACG CGCACAGTGC ATCCATCAGT GCAACTCCGC ACCACGCCAT CATCAGCAAC ACACCCCCTA AAAAAAAAAA

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                2570       2580       2590       2600       2610       2620       2630       2640
```

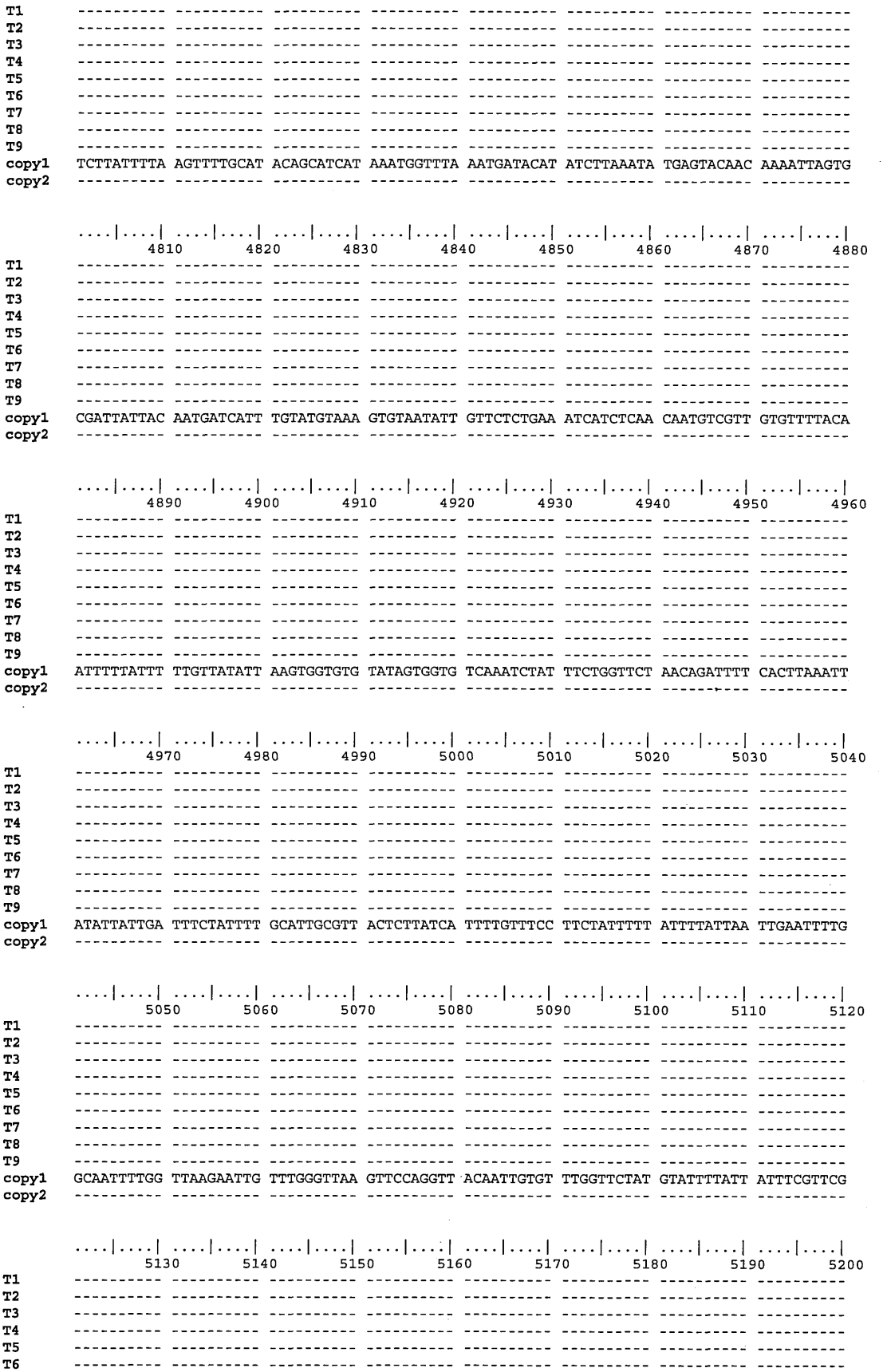

T1
T2
T3
T4
T5
T6
T7
T8
T9
copy1 CCCAGTATTC CTTCCATGAC ACGTCACTCA TCTCATCTCA TCGCATTTTT GCATTCCTGC ATCGCTTGCA TCCTCCGCCT
copy2 CCCAGTATTC CTTCCATGAC ACGTCACTCA TCTCATCTCA TCGCATTTTT GCATTCCTGC ATCGCTTGCA TCCTCCGCCT
2650 2660 2670 2680 2690 2700 2710 2720
copy1 AACCGCACAA AATCACACAC TAACTATTTT CGATACCCGA ATCAAATATT AACTATATTT GAATCCTTAA CTTTGTCAAT
copy2 AACCGCACAA AATCACACAC TAACTATTTT CGATACCCGA ATCAAATATT AACTATATTT GAATCCTTAA CTTTGTCAAT
2730 2740 2750 2760 2770 2780 2790 2800
copy1 CGAAAGAATC TTTTGCATCT GAAAAGAAAG AGCACTAAAG TCAAAGATAG ATCGGAATTA AAACGAAGTT GCAACAAGAA
copy2 CGAAAGAATC TTTTGCATCT GAAAAGAAAG AGCACTAAAG TCAAAGATAG ATCGGAATTA AAACGAAGTT GCAACAAGAA
2810 2820 2830 2840 2850 2860 2870 2880
copy1 GTGCATAATC AGATATTTCG ACTGAAATAA GCTATTTAAA ATATTTAAAA AATATTTCAA ACATATCTTG TAGTGTGTGC
copy2 GTGCATAATC AGATATTTCG ACTGAAATAA GCTATTTAAA ATATTTAAAA AATATTTCAA ACATATCTTG TAGTGTGTGC
2890 2900 2910 2920 2930 2940 2950 2960
copy1 TTTTGTGTGT GTTGCTTTGT AAAATCCTTT GTACTTTAGT TGAAGTTGAT TTTGTTCCGC TGTTCCTTAG AAATGTATTT
copy2 TTTTGTGTGT GTTGCTTTGT AAAATCCTTT GTACTTTAGT TGAAGTTGAT TTTGTTCCGC TGTTCCTTAG AAATGTATTT
2970 2980 2990 3000 3010 3020 3030 3040
T3 -----ATTGT TGCCTAGAGT TTTCATTTGC TGACGCTGCT GTCGCCTGTC

```
T7      ---------- ---------- ---------- -----ATTGT TGCCTAGAGT TTTCATTTGC TGACGCTGCT GTCGCCTGTC
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   ATGATTAACC ATAGCTACAA ACGTTGCTCC TCCAGATTGT TGCCTAGAGT TTTCATTTGC TGACGCTGCT GTCGCCTGTC
copy2   ATGATTAACC ATAGCTACAA ACGTTGCTCC TCCAGATTGT TGCCTAGAGT TTTCATTTGC TGACGCTGCT GTCGCCTGTC

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
              3050       3060       3070       3080       3090       3100       3110       3120
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      GCCCGTTGCC TCCTGCCTGC CTGCTCTGCT CCTGCCTCCT GGCATCCCGC ATCCTGCATC CTGGCACCTG CCAACCGACA
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      GCCCGTTGCC TCCTGCCTGC CTGCTCTGCT CCTGCCTCCT GGCATCCCGC ATCCTGCATC CTGGCACCTG CCAACCGACA
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GCCCGTTGCC TCCTGCCTGC CTGCTCTGCT CCTGCCTCCT GGCATCCCGC ATCCTGCATC CTGGCACCTG CCAACCGACA
copy2   GCCCGTTGCC TCCTGCCTGC CTGCTCTGCT CCTGCCTCCT GGCATCCCGC ATCCTGCATC CTGGCACCTG CCAACCGACA

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
              3130       3140       3150       3160       3170       3180       3190       3200
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      CCTGCTCCTC CCGACCCAAC CGTAGCATCT AAC------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      CCTGCTCCTC CCGACCCAAC C--------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CCTGCTCCTC CCGACCCAAC CGTAGCATCT AACCCGCCCG CTAATAATCT TCACTTCCGT CAGTGATTGT CCGTTCTGCT
copy2   CCTGCTCCTC CCGACCCAAC CGTAGCATCT AACCCGCCCG CTAATAATCT TCACTTCCGT CAGTGATTGT CCGTTCTGCT

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
              3210       3220       3230       3240       3250       3260       3270       3280
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GTCGCTTACT CTACCAACTC TATAGTAGGT CAGGGGTTAG CCAATATATA TGCATATATA TCTTTCTAGA AGATTTGAAG
copy2   GTCGCTTACT CTACCAACTC TATAGTAGGT CAGGGGTTAG CCAATATATA TGCATATATA TCTTTCTAGA AGATTTGAAG

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
              3290       3300       3310       3320       3330       3340       3350       3360
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   ATAGCGATTT CACAAGATCT GATAAAGTGT TTGGATGATA CATCACATCA ATTTGTCATT TGTAATCAGC AAAAGGCGCT
copy2   ATAGCGATTT CACAAGATCT GATAAAGTGT TTGGATGATA CATCACATCA ATTTGTCATT TGTAATCAGC AAAAGGCGCT

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
              3370       3380       3390       3400       3410       3420       3430       3440
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AATCAAAAAT TTCACCACTT GTACGCTTGG GCTATTAAAA CCCTAGAAGT TTCTGGAGGT TGAATTAAAC AATTTAATTG
copy2   AATCAAAAAT TTCACCACTT GTACGCTTGG GCTATTAAAA CCCTAGAAGT TTCTGGAGGT TGAATTAAAC AATTTAATTG
```

```
          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  3450       3460       3470       3480       3490       3500       3510       3520
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     TTTGAAAGAA CAAATGTAGT TTTTACTTGC ATAGATATTA AACTCACTCA ACATTTTGGT CATACGTCGT TTTAGACGAA
copy2     TTTGAAAGAA CAAATGTAGT TTTTACTTGC ATAGATATTA AACTCACTCA ACATTTTGGT CATACGTCGT TTTAGACGAA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  3530       3540       3550       3560       3570       3580       3590       3600
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     TCACACCTGT AGTTAGGAAG CCCCACATGG AGAGCCCACA GCTGTACTGC AGATAGGGTT ACAGATAGGG ATACGCTGAT
copy2     TCACACCTGT AGTTAGGAAG CCCCACATGG AGAGCCCACA GCTGTACTGC AGATAGGGTT ACAGATAGGG ATACGCTGAT

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  3610       3620       3630       3640       3650       3660       3670       3680
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     ATAGATACGC CTGGTAATTG TTTATTTAAA GTCGAGGCGT GTGCGATAAC AGACCCAATA AGTGCATTTG TGGGTGCAAA
copy2     ATAGATACGC CTGGTAATTG TTTATTTAAA GTCGAGGCGT GTGCGATAAC AGACCCAATA AGTGCATTTG TGGGTGCAAA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  3690       3700       3710       3720       3730       3740       3750       3760
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     ATGTTTGTCT TCAATTACAA ATTCAAGACT CGAATTTATG AAATGCCAGC AGAACTCAAA ACACTTTCAT GCACGTTTCG
copy2     ATGTTTGTCT TCAATTACAA ATTCAAGACT CGAATTTATG AAATGCCAGC AGAACTCAAA ACACTTTCAT GCACGTTTCG

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  3770       3780       3790       3800       3810       3820       3830       3840
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     AGTTTTTAGC TGTTTGTCTA TTTGTTTAAT TGCCGGCAAT GGTGACGTAA ACATAATTAT AAACAGATTA TATCAGAGCA
copy2     AGTTTTTAGC TGTTTGTCTA TTTGTTTAAT TGCCGGCAAT GGTGACGTAA ACATAATTAT AAACAGATTA TATCAGAGCA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  3850       3860       3870       3880       3890       3900       3910       3920
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

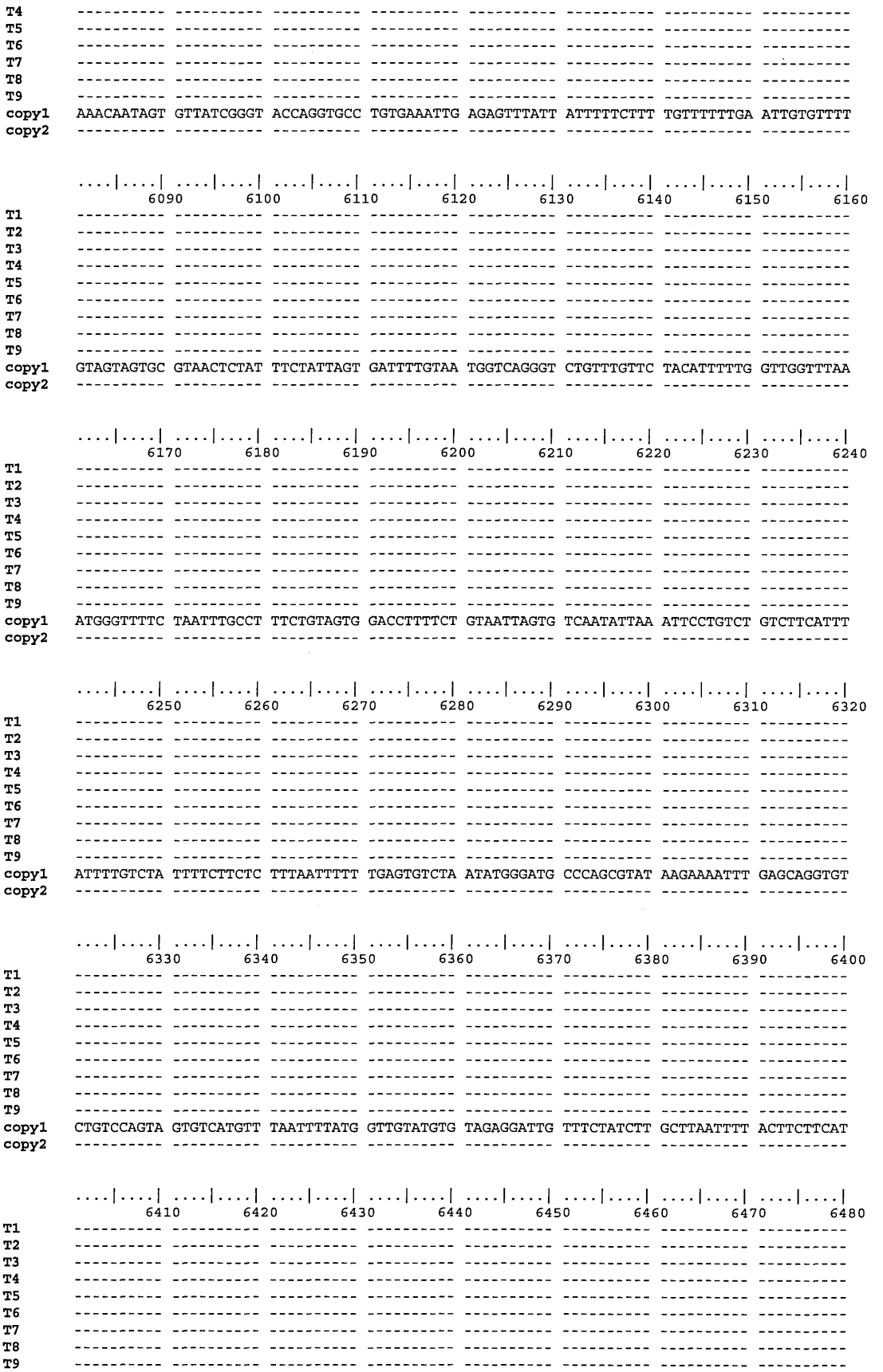
```
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGTTTATAAA ATAGTCAAAT TACCAGCTGC AAATGGGCTA ATAACTTGTT ATTACTTCGT CATGCTTAGT TACATATGGA
copy2   TGTTTATAAA ATAGTCAAAT TACCAGCTGC AAATGGGCTA ATAACTTGTT ATTACTTCGT CATGCTTAGA TACATATGGA

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                3930       3940       3950       3960       3970       3980       3990       4000
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AAATAGGTAT TAGTAGTAGT ATTATTATTA TTATTTCATT CAATAACAGC TTGTAGTCAT ATCGGAAATA TTTAAACAAT
copy2   AAATAGGTAT TAGTAGTAGT ATTATTATTA TTATTTCATT CAATAACAGC TTGTAGTCAT ATCGGAAATA TTTAAACAAT

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4010       4020       4030       4040       4050       4060       4070       4080
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TTGAAACTGG CGGCGTTTGT GGCGAAATAT TTGGAGTCTT TCATGGCCCT AAAAATTATA GAGTTAATAG CATTGGGAGA
copy2   TTGAAACTGG CGGCGTTTGT GGCGAAATAT TTGGAGTCTT TCATGGCCCT AAAAATTATA GAGTTAATAG CATTGGGAGA

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4090       4100       4110       4120       4130       4140       4150       4160
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGCATCGTTT CTTGCACCAT TTCTCGCCTC TCTCTCTCTC TCCGTCTCTT TTTGAATATA TATAATTAAA AATTCGCACC
copy2   TGCATCGTTT CTTGCACCAT TTCTCGCCTC TCTCTCTCTC TCCGTCTCTT TTTGAATATA TATAATTAAA AATTCGCACC

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4170       4180       4190       4200       4210       4220       4230       4240
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGCCTCATCC TACATATATA TACATATATA TATATAGTAA AATTGTTGTT TTATTTATGA GTCGAACTAA TGTCCCGTCA
copy2   TGCCTCATCC TACATATATA TACATATATA TATATA---- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4250       4260       4270       4280       4290       4300       4310       4320
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
copy1   GTTGACTAAG GACAACGCCA AGAAATAAAA AATAGTTTTT AGGAGAAATA GTTTCACGAC TTTATTCTTT GGATCTCAGC
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4330       4340       4350       4360       4370       4380       4390       4400
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TTAAGACTAA AAATCAAATG CAAATTGGAT TGAGGAGAAG CCTCAGTATT TGAACAATAT TTGTATTTCG GGAGAGCCTC
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4410       4420       4430       4440       4450       4460       4470       4480
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GCTCTTGGGT TCTTAAGATT AGGTAGCGTT GAAAGAGGGC AGTCAAATCT GCTTAGGTCA GGCTTTAGGA TGTTTACAAG
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4490       4500       4510       4520       4530       4540       4550       4560
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AACGGCTGCG CTGCCAAAGA TAAACGAATG TTGCGCAGCT GGGATACGTT ATGGAAAATT ACTAGAGTGC ATTACTGATT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4570       4580       4590       4600       4610       4620       4630       4640
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TCTTTGAAAT AATGGAAGTG GCTGGTTTGG ACCACCCAAA TACGTCACTC CCCTTCCCTT AAAGAGAAGC CTATACTTGG
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4650       4660       4670       4680       4690       4700       4710       4720
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GCTGGGATTG ATGGATATAG TGTGGGAAAT GGAGTTTCTT CTATTTCTGT TTGTTGTGGT GTGTTTTGAT TTTGATTTTT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                4730       4740       4750       4760       4770       4780       4790       4800
```


T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 TCTTATTTTA AGTTTTGCAT ACAGCATCAT AAATGGTTTA AATGATACAT ATCTTAAATA TGAGTACAAC AAAATTAGTG
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
4810 4820 4830 4840 4850 4860 4870 4880
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 CGATTATTAC AATGATCATT TGTATGTAAA GTGTAATATT GTTCTCTGAA ATCATCTCAA CAATGTCGTT GTGTTTTACA
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
4890 4900 4910 4920 4930 4940 4950 4960
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 ATTTTTATTT TTGTTATATT AAGTGGTGTG TATAGTGGTG TCAAATCTAT TTCTGGTTCT AACAGATTTT CACTTAAATT
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
4970 4980 4990 5000 5010 5020 5030 5040
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 ATATTATTGA TTTCTATTTT GCATTGCGTT ACTCTTATCA TTTTGTTTCC TTCTATTTTT ATTTTATTAA TTGAATTTTG
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
5050 5060 5070 5080 5090 5100 5110 5120
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 GCAATTTTGG TTAAGAATTG TTTGGGTTAA GTTCCAGGTT ACAATTGTGT TTGGTTCTAT GTATTTTATT ATTTCGTTCG
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
5130 5140 5150 5160 5170 5180 5190 5200
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

```
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGTGTACTGG CTTAAAATTA CAAATTGGAT TTAAGTGTTT AATAATATTG GTGACACATT CATTTTTTAC TTCTTTATTT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5210       5220       5230       5240       5250       5260       5270       5280
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TCGGTATTAT AAACTTTATT TTCTTTTTCA AAATATGATT GTGTGTCTGT GTAATCTAGC TGATAGCCGT TGGAGTCTGG
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5290       5300       5310       5320       5330       5340       5350       5360
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GTAAGGAATT ATTTTAATTG TGCTTATTAG TGAAAGGTAA ATGGGTATGT GGGATATGAT TAATACTTGG TTATCATTGT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5370       5380       5390       5400       5410       5420       5430       5440
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AGTTAATCCA AGTGGATGTT TTAATGTTCA AAATGTTTTG GCTGTTCACA TTTTCAAGTT TGTCGTAGTT TAGTAATTTG
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5450       5460       5470       5480       5490       5500       5510       5520
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GGGTTAAACA GTCCAAGTCT GGAAAGCTGC ATAGCCATTT CCAAATCTTC TATGTATTCC GTAAACTGCA TTAGTTCGAA
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5530       5540       5550       5560       5570       5580       5590       5600
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CAAAAGACTG TCAATGATGG TGTGGTTATT TTCATAATTC TGTATTTTTT GCATTCCGTC ATTTATCAAT CGTATGGCGT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5610       5620       5630       5640       5650       5660       5670       5680
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     CGTTGAGTTC ATGTAATTTT ACTGAGTTAT GGACGTTGAT TTCAAGTTTT TTCTGTATTT CGACTCGATC ATTTTCATCT
copy2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5690       5700       5710       5720       5730       5740       5750       5760
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     AGTGTTCCAA ATAGGTATTT AAAACCTGAA CCCACAATGT TAATAAGACC CCGTTTGTTT CTATGTTTCA GGGTTATTCC
copy2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5770       5780       5790       5800       5810       5820       5830       5840
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     GGCTAGTTCT CTTTTTATTT TATTGTATAA AAAATTGATC TGGGGTGCAT GGTAATCGGT TCGTAGTCTT TGCTCAAAAT
copy2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5850       5860       5870       5880       5890       5900       5910       5920
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     AGTCAACCAC GATGTCTATT TCAGTTAAAT TGATTTTTAA GGAGTGATGT TCAAAGGAGG ATGGTATCTG TACTGGCTTA
copy2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5930       5940       5950       5960       5970       5980       5990       6000
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     TCAGAAAAAA GGAGATATCC GTGGTTGGTG TCAATATTAT TAATTTGAAT TTGTTGTCCA TGAACAGCTG TGATCAGTAT
copy2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  6010       6020       6030       6040       6050       6060       6070       6080
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```


T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 AAACAATAGT GTTATCGGGT ACCAGGTGCC TGTGAAATTG AGAGTTTATT ATTTTTCTTT TGTTTTTTGA ATTGTGTTTT
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
6090 6100 6110 6120 6130 6140 6150 6160
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 GTAGTAGTGC GTAACTCTAT TTCTATTAGT GATTTTGTAA TGGTCAGGGT CTGTTTGTTC TACATTTTTG GTTGGTTTAA
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
6170 6180 6190 6200 6210 6220 6230 6240
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 ATGGGTTTTC TAATTTGCCT TTCTGTAGTG GACCTTTTCT GTAATTAGTG TCAATATTAA ATTCCTGTCT GTCTTCATTT
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
6250 6260 6270 6280 6290 6300 6310 6320
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 ATTTTGTCTA TTTTCTTCTC TTTAATTTTT TGAGTGTCTA ATATGGGATG CCCAGCGTAT AAGAAAATTT GAGCAGGTGT
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
6330 6340 6350 6360 6370 6380 6390 6400
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1 CTGTCCAGTA GTGTCATGTT TAATTTTATG GTTGTATGTG TAGAGGATTG TTTCTATCTT GCTTAATTTT ACTTCTTCAT
copy2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
6410 6420 6430 6440 6450 6460 6470 6480
T1 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

```
copy1   CATCAGATGA ATTGATTATA CGAATTTTTT CATTTATTGT TTTGTGTAAT CTTTCGACGT CTGCTACCCC GTTTTTTGCT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6490       6500       6510       6520       6530       6540       6550       6560
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GTATTGAGCT GTAATCCTAC CGACTGCGCC AGTAAAATTG TTGTTTTATT TATGAGTCGA ACTAATGTCC CGTCAGTTGA
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6570       6580       6590       6600       6610       6620       6630       6640
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CTAAGGACAA CGCCAAGAAA TAAAAAATAG TTTTTAGGAG AAATAGTTTC ACGACTTTAT TCTTTGGATC TCAGCTTAAG
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6650       6660       6670       6680       6690       6700       6710       6720
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   ACTAAAAATC AAATGCAAAT TGGATTGAGG AGAAGCCTCA GTATTTGAAC AATATTTGTA TTTCGGGAGA GCCTCGCTCT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6730       6740       6750       6760       6770       6780       6790       6800
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TGGGTTCTTA AGATTAGGTA GCGTTGAAAG AGGGCAGTCA AATCTGCTTA GGTCAGGCTT TAGGATGTTT ACAAGAACGG
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6810       6820       6830       6840       6850       6860       6870       6880
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CTGCGCTGCC AAAGATAAAC GAATGTTGCG CAGCTGGGAT ACGTTATGGA AAATTACTAG AGTGCATTAC TGATTTCTTT
copy2   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6890       6900       6910       6920       6930       6940       6950       6960
```

```
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GAAATAATGG AAGTGGCTGG TTTGGACCAC CCAAATACGT CACTATATAT ATATATATAT GCACGTATAT ATCTATGGAC
copy2   ---------- ---------- ---------- ---------- -------TAT ATATA----T GCACGTATAT ATCTATGGAC

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                6970       6980       6990       7000       7010       7020       7030       7040
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GAGGATCGTA AACCATACCT AGACATATTG AATCGCTTTA CCGGCTAAGC AGCTAATCTT TAAGGGACGC ACCTAATGTA
copy2   GAGGATCGTA AACCATACCT AGACATATTG AATCGCTTTA CCGGCTAAGC AGCTAATCTT TAAGGGACGC ACCTAATGTA

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                7050       7060       7070       7080       7090       7100       7110       7120
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- --CATATAGC ATCGGTCTGT TGCGATCCGA CTATATGGCA CACGTGTCCG
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- -GCATATAGC ATCGGTCTGT TGCGATCCGA CTATATGGCA CACGTGTCCG
T8      ---------- ---------- ---------- -GCATATAGC ATCGGTCTGT TGCGATCCGA CTATATGGCA CACGTGTCCG
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TATATATCTT GTATAACGAT CATAATTGCA GGCATATAGC ATCGGTCTGT TGCGATCCGA CTATATGGCA CACGTGTCCG
copy2   TATATATCTT GTATAACGAT CATAATTGCA GGCATATAGC ATCGGTCTGT TGCGATCCGA CTATATGGCA CACGTGTCCG

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                7130       7140       7150       7160       7170       7180       7190       7200
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      AGGGCTGTGC GATAAAGCAG GTCGAGATCA ACACGGTGGC CTCCAGTTTC GCTGGCATAG CCACCCAATT GGTGCCGCAG
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      AGGGCTGTGC GATAAAGCAG GTCGAGATCA ACACGGTGGC CTCCAGTTTC GCTGGCATAG CCACCCAATT GGTGCCGCAG
T8      AGGGCTGTGC GATAAAGCAG GTCGAGATCA ACACGGTGGC CTCCAGTTTC GCTGGCATAG CCACCCAATT GGTGCCGCAG
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AGGGCTGTGC GATAAAGCAG GTCGAGATCA ACACGGTGGC CTCCAGTTTC GCTGGCATAG CCACCCAATT GGTGCCGCAG
copy2   AGGGCTGTGC GATAAAGCAG GTCGAGATCA ACACGGTGGC CTCCAGTTTC GCTGGCATAG CCACCCAATT GGTGCCGCAG

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                7210       7220       7230       7240       7250       7260       7270       7280
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      CAGAA----- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      CAGAA----- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      CAGAA----- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CAGAAGTATT AAGCGCTCTT CTCTCTATCT CTTTCTACAT ATAATATATA TAAATATATA TCTATCAAAC TACATATATT
copy2   CAGAAGTATT AAGCGCTCTT CTCTCTATCT CTTTCTACAT ATAATATATA TAAATATATA TCTATCAAAC TACATATATT

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                7290       7300       7310       7320       7330       7340       7350       7360
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

```
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CCTTTTACCA AACGCACGAA TGCACAATGC ACACACCTAA AAGCAAAAGC ACACTACACA CCTGTCCTAA TATCCTTCCA
copy2   CCTTTTACCA AACGCACGAA TGCACAATGC ACACACCTAA AAGCAAAAGC ACACTACACA CCTGTCCTAA TATCCTTCCA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7370       7380       7390       7400       7410       7420       7430       7440
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   CCCAACTCCT TCTCCTTCTC ACTGATCACT TATCCACCTA CAGAGGTGCC ATCTCTATCG TTTGTTGTCT CTTTCTGCAT
copy2   CCCAACTCCT TCTCCTTCTC ACTGATCACT TATCCACCTA CAGAGGTGCC ATCTCTATCG TTTGTTGTCT CTTTCTGCAT

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7450       7460       7470       7480       7490       7500       7510       7520
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   GTGTTTTTTT TTTTTTTAAT TGGTTATGGC TCTATCCGTC TCTTTCCACA CGCTGTAGCT TCATAACCCC ACTTCCACAA
copy2   GTGTTTTTTT TTTTTTTAAT TGGTTATGGC TCTATCCGTC TCTTTCCACA CGCTGTAGCT TCATAACCCC ACTTCCACAA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7530       7540       7550       7560       7570       7580       7590       7600
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   TTTCCCTAAC CATAAACATA CATATGTATA TCCATCCGCC CATTTGAGAG GTTTTGGTCT CAGTAATTAA CCCAAATTTC
copy2   TTTCCCTAAC CATAAACATA CATATGTATA TCCATCCGCC CATTTGAGAG GTTTTGGTCT CAGTAATTAA CCCAAATTTC

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7610       7620       7630       7640       7650       7660       7670       7680
T1      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1   AACCCCCCCC CCC--ATTAC CACCGAACAT TTGTTGTTGT TTGTTGCTGT TGTTGGTCGC TGAAAGACTG CGGAATTATG
copy2   AACCCCCCCC CCCCCATTAC CACCGAACAT TTGTTGTTGT TTGTTGCTGT TGTTGGTCGC TGAAAGACTG CGGAATTATG

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7690       7700       7710       7720       7730       7740       7750       7760
T1      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
T2      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
T3      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---.......
T4      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
T5      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
T6      ---------- ---------- ---------- ~~~~~~~~~~ ~~~~~~~~~~ ---ATTCGTA CTCAGCGAGC TGGGACATGC
T7      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
T8      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
T9      ---------- ---------- ---------- ---------- ---------- ---ATTCGTA CTCAGCGAGC TGGGACATGC
copy1   CTAATGTCTG TCTCTTTCTG TTCCGGCACT ACCCGTCTCT TTCGCCCGCG CAGATTCGTA CTCAGCGAGC TGGGACATGC
copy2   CTAATGTCTG TCTCTTTCTG TTCCGGCACT ACCCGTCTCT TTCGCCCGCG CAGATTCGTA CTCAGCGAGC TGGGACATGC
```

```
          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7770       7780       7790       7800       7810       7820       7830       7840
T1        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T2        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T3        .......... .......... .......... .......... .......... .......... .......... ..........
T4        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T5        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T6        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T7        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T8        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
T9        GGATAAGCTG CACAAT---- ---------- ---------- ---------- ---------- ---------- ----------
copy1     GGATAAGCTG CACAATGTGA GTATATTAAA CTAAGATATA TAATCTATAT ATATATAAAC ATATATTCAT ATATATAATA
copy2     GGATAAGCTG CACAATGTGA GTATATTAAA CTAAGATATA TAATCTATAT ATATATAAAC ATATATTCAT ATATATAATA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7850       7860       7870       7880       7890       7900       7910       7920
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T3        .......... .......... .......... .......... .......... .......... .......... ..........
T4        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T5        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T6        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T7        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T8        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
T9        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
copy1     ATTATATGAT CTGTGTTTCC TGTACAATAA ACCATGAGTA ACAAGACAAC TTAAATGATA ATGTTCGTTA TTCCGCTGCA
copy2     ATTATATGAT CTGTGTTTCC TGTACAATAA ACCATGAGTA ACAAGACAAC TTAAATGATA ATGTTCGTTA TTCCGCTGCA

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  7930       7940       7950       7960       7970       7980       7990       8000
T1        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T2        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T3        .......... .......... .......... .......... .......... .......... .......... ..........
T4        -ATGCCGGAT AACAATGCAC TGGCCGGACT CCGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T5        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T6        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T7        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T8        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
T9        -ATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
copy1     GATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG
copy2     GATGCCGGAT AACAATGCAC TGGCCGGACT CTGCGATGGC ATGGTGAAGG CGTGGGACAT CTATGCAAAG CCGCAGGCCG

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  8010       8020       8030       8040       8050       8060       8070       8080
T1        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T2        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T3        .......... .......... .......... .......... .......... .......... .......... ..........
T4        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T5        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T6        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T7        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T8        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
T9        TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
copy1     TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT
copy2     TCATTCTGTT CATCATCGAG GATGTGTCGT ACAACATTTG CGACCAGCGG TTCCACGAGT TCTACATTCG CGAGACATAT

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  8090       8100       8110       8120       8130       8140       8150       8160
T1        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T2        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T3        .......... .......... .......... .......... .......... .......... .......... ..........
T4        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T5        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T6        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T7        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T8        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
T9        CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
copy1     CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT
copy2     CCGCACATCA AGGTGCTGCG CCGCACCCTC ACCGAGGTCC ATCGGGAGGG CAAACTGGGC CAGTCCAAGG AGTTGCTCCT

          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  8170       8180       8190       8200       8210       8220       8230       8240
T1        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T2        ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T3        .......... .......... .......... .......... .......... .......... .......... ..........
```

```
T4      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T5      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T6      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T7      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T8      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
T9      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------TGGA
copy1   GTAAGTGTAC ATTAATACCT CTATTATCAT TCATCATTGA ACTGTGTGCT CAACTCACAC GCTTGCCGCA CCACAGTGGA
copy2   GTAAGTGTAC ATTAATACCT CTATTATCAT TCATCATTGA ACTGTGTGCT CAACTCACAC GCTTGCCGCA CCACAGTGGA

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                      8250        8260        8270        8280        8290        8300        8310        8320
T1      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T2      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T5      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T6      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T7      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T8      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
T9      TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
copy1   TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG
copy2   TCCCAAGAGG TGGCCGTCAT CTACTTCCGA GCTGGCTACG AGCCCGGACA CTATCACTCG CAGGCGGATT GGGATGCCCG

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                      8330        8340        8350        8360        8370        8380        8390        8400
T1      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T2      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T5      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T6      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T7      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T8      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
T9      CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
copy1   CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG
copy2   CTATCTGATG GAGACCTCGC TGGCCATCAA GTGCCCGTCG ATTCACTACC ATTTGGCGGG CACCAAGAAG GTGCAGCAGG

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                      8410        8420        8430        8440        8450        8460        8470        8480
T1      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T2      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T5      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T6      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T7      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T8      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
T9      CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
copy1   CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT
copy2   CATTGGCGCA GCCGGCGGTG CTCGAGCGTT TCATCAACGA TCCGGAGGAG ATCAAGGCGG TGGGCAAGAT CTTCACGGGT

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                      8490        8500        8510        8520        8530        8540        8550        8560
T1      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T2      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T5      CTCTACTCGC TGGACGATAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T6      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T7      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T8      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
T9      CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
copy1   CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA
copy2   CTCTACTCGC TGGACGACAA CGAGGCGGGC AATGCCAGCT ACGAGATGGC GCTGCGCACT CCGGAGAGAT TCGTGCTGAA

                .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                      8570        8580        8590        8600        8610        8620        8630        8640
T1      GCCGCAGCGC GAGGGAGGCG GCAATAAT-- ---------- ---------- ---------- ---------- ----------
T2      GCCGCAGCGC GAGGGAGGCG GCAATAAT-- ---------- ---------- ---------- ---------- ----------
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAA ATGCACTAGA G--------- ----------
T5      GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAN ATGCACTAAA G--------- ----------
T6      GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAG ATGCACTAAA G--------- ----------
T7      GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAG ATGCACTAAA G--------- ----------
T8      GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAG ATGCACTAAA G--------- ----------
T9      GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAG ATGCACTAAA G--------- ----------
```

```
copy1   GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAG ATGCACTAAA GGTGAGTTAT TCCACTTTGA
copy2   GCCGCAGCGC GAGGGAGGCG GCAATAATGT CTACGGCGTG GATATACCAG ATGCACTAAA GGTGAGTTAT TCCACTTTGA

                 8650       8660       8670       8680       8690       8700       8710       8720
T1      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T2      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T5      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T6      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T7      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T8      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
T9      ---------- ---------- ---------- ---------- ---------- --------CG CATGTCACGC GTGGAGCGCT
copy1   TGATTATTTG AATTGTGAAC AGTTTTGTAA CAAATCCTTC TCCTTCTTTA TCCATCAGCG CATGTCACGC GTGGAGCGCT
copy2   TGATTATTTG AATTGTGAAC AGTTTTGTAA CAAATCCTTC TCCTTCTTTA TCCATCAGCG CATGTCACGC GTGGAGCGCT

                 8730       8740       8750       8760       8770       8780       8790       8800
T1      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T2      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T5      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T6      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T7      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T8      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
T9      CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
copy1   CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC
copy2   CCGCCTGGAT TCTGATGGAT CTGATTCATC CGCCGCTGAC CAAGGGCTAT ATGGTGCGGC CCGGTGGCGA TATGCCGCCC

                 8810       8820       8830       8840       8850       8860       8870       8880
T1      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T2      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T5      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T6      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T7      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T8      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
T9      CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
copy1   CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA
copy2   CAGATCGTGG ACATGGTCTC CGAGCTGGGC ATCTTCGGTG TGGTGATCGG AGATGCGGAG CACATTGTGC ACAACTACCA

                 8890       8900       8910       8920       8930       8940       8950       8960
T1      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T2      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T5      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T6      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T7      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T8      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
T9      GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
copy1   GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC
copy2   GGCGGGACAC ATGCTGCGCA CAAAGCTCTC AACCGCCAAC GAAGGAGGCG TGGCAGCCGG TCTGGGCGCC CTGGACAGTC

                 8970       8980       8990       9000       9010       9020       9030       9040
T1      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA GG........ .......... .......... ..........
T2      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA GG-------- .......... .......... ..........
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA GGATCCTTTA ACGAGATAAA GGCTTCGAAC GCCGGCATTC
T5      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA GGATCCTTTA ACGAGATAAA GGCTTCGAAC GCCGGCATTC
T6      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA G......... .......... .......... ..........
T7      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA G......... .......... .......... ..........
T8      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA G......... .......... .......... ..........
T9      CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA G......... .......... .......... ..........
copy1   CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA GGATCCTTTA ACGAGATAAA GGCTTCGAAC GCCGGCATTC
copy2   CATATCTGAT CGATAGCGAT GACGAGGATG AGCAGAAGTA GGATCCTTTA ACGAGATAAA GGCTTCGAAC GCCGGCATTC

                 9050       9060       9070       9080       9090       9100       9110       9120
```

```
T1      .......... .......... .......... .......... .......... .......... .......... ..........
T2      .......... .......... .......... .......... .......... .......... .......... ..........
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CACCTGTTGC CTCAGTCTTG TATAAATGAT TATTTTTTGT AAAGCGTATT GTAAATTTTT TGCACTATTT TTTCCCCGTC
T5      CACCTGTTGC CTCAGTCTTG TATAAATGAT TATTTTTTGT AAAGCGTATT GTAAATTTTT TGCACTATTT TTTCCCCGTC
T6      .......... .......... .......... .......... .......... .......... .......... ..........
T7      .......... .......... .......... .......... .......... .......... .......... ..........
T8      .......... .......... .......... .......... .......... .......... .......... ..........
T9      .......... .......... .......... .......... .......... .......... .......... ..........
copy1   CACCTGTTGC CTCAGTCTTG TATAAATGAT TATTTTT-GT AAAGCGTATT GTAAATTTTT TGCACTATTT TTTCCCCGTC
copy2   CACCTGTTGC CTCAGTCTTG TATAAATGAT TATTTTT-GT AAAGCGTATT GTAAATTTTT TGCACTATTT TTTCCCCGTC

        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                9130       9140       9150       9160       9170       9180       9190       9200
T1      .......... .......... .......... .......... .......... .......... .......... ..........
T2      .......... .......... .......... .......... .......... .......... .......... ..........
T3      .......... .......... .......... .......... .......... .......... .......... ..........
T4      CCAATGTCAA TAAATAAAAT GTTAA-CTAG AAAAAAAAAA AAAAAAAACT CGAGGGGGGG CCCGGCCCCA ATTCCCCCTT
T5      CCAATGTCAA TAAATAAAAT GTTAA-CTAG AAAAAAAAAA AAAAAAA-CT CNAGGGGGGG CCCGGNCCCA ATTCCCCCTT
T6      .......... .......... .......... .......... .......... .......... .......... ..........
T7      .......... .......... .......... .......... .......... .......... .......... ..........
T8      .......... .......... .......... .......... .......... .......... .......... ..........
T9      .......... .......... .......... .......... .......... .......... .......... ..........
copy1   CCAATGTCAA TAAATAAAAT GTTAAACTAG AAATTCTGGT CAATATGCAA ACTATGCAAA ATATGCCAGG ..........
copy2   CCAATGTCAA TAAATAAAAT GTTAAACTAG AAATTCTGGT CAATATGCAA ACTGGAT--- ---------- ..........

        ....|....| ..
                9210
T1      .......... ..
T2      .......... ..
T3      .......... ..
T4      TATGGAGTAT --
T5      TANGGAGTNT GT
T6      .......... ..
T7      .......... ..
T8      .......... ..
T9      .......... ..
copy1   .......... ..
copy2   .......... ..
```

9.3 Appendix C: Sequence of human GS cDNA cloned into pUAST

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
        10         20         30         40         50         60         70         80         90
GGCGAACTAG TGTTGGGATG GCCACCAACT GGGGGAGCCT CTTGCAGGAT AAACAGCAGC TAGAGGAGCT GGCACGGCAG GCCGTGGACC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       100        110        120        130        140        150        160        170        180
GGGCCCTGGC TGAGGGAGTA TTGCTGAGGA CCTCACAGGA GCCCACTTCC TCGGAGGTGG TGAGCTATGC CCCATTCACG CTCTTCCCCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       190        200        210        220        230        240        250        260        270
CACTGGTCCC CAGTGCCCTG CTGGAGCAAG CCTATGCTGT GCAGATGGAC TTCAACCTGC TAGTGGATGC TGTCAGCCAG AACGCTGCCT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       280        290        300        310        320        330        340        350        360
TCCTGGAGCA AACTCTTTCC AGCACCATCA AACAGGATGA CTTTACCGCT CGTCTCTTTG ACATCCACAA GCAAGTCCTA AAAGAGGGCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       370        380        390        400        410        420        430        440        450
TTGCCCAGAC TGTGTTCCTG GGCCTGAATC GCTCAGACTA CATGTTCCAG CGCAGCGCAG ATGGCTCCCC AGCCCTGAAA CAGATCGAAA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       460        470        480        490        500        510        520        530        540
TCAACACCAT CTCTGCCAGC TTTGGGGGCC TGGCCTCCCG GACCCCAGCT GTGCACCGAC ATGTTCTCAG TGTCCTGAGT AAGACCAAAG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       550        560        570        580        590        600        610        620        630
AAGCTGGCAA GATCCTCTCT AATAATCCCA GCAAGGGACT GGCCCTGGGA ATTGCCAAAG CCTGGGAGCT CTACGGCTCA CCCAATGCTC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       640        650        660        670        680        690        700        710        720
TGGTGCTACT GATTGCTCAA GAGAAGGAAA GAAACATATT TGACCAGCGT GCCATAGAGA ATGAGCTACT GGCCAGGAAC ATCCATGTGA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       730        740        750        760        770        780        790        800        810
TCCGACGAAC ATTTGAAGAT ATCTCTGAAA AGGGGTCTCT GGACCAAGAC CGAAGGCTGT TTGTGGATGG CCAGGAAATT GCTGTGGTTT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       820        830        840        850        860        870        880        890        900
ACTTCCGGGA TGGCTACATG CCTCGTCAGT ACAGTCTACA GAATTGGGAA GCACGTCTAC TGCTGGAGAG GTCACATGCT GCCAAGTGCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
       910        920        930        940        950        960        970        980        990
CAGACATTGC CACCCAGCTG GCTGGGACTA AGAAGGTGCA GCAGGAGCTA AGCAGGCCGG GCATGCTGGA GATGTTGCTC CCTGGCCAGC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1000       1010       1020       1030       1040       1050       1060       1070       1080
CTGAGGCTGT GGCCCGCCTC CGCGCCACCT TTGCTGGCCT CTACTCACTG GATGTGGGTG AAGAAGGGGA CCAGGCCATC GCCGAGGCCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1090       1100       1110       1120       1130       1140       1150       1160       1170
TTGCTGCCCC TAGCCGGTTT GTGCTAAAGC CCCAGAGAGA GGGTGGAGGT AACAACCTAT ATGGGGAGGA AATGGTACAG GCCCTGAAAC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1180       1190       1200       1210       1220       1230       1240       1250       1260
AGCTGAAGGA CAGTGAGGAG AGGGCCTCCT ACATCCTCAT GGAGAAGATC GAACCTGAGC CTTTTGAGAA TTGCCTGCTA CGGCCTGGCA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1270       1280       1290       1300       1310       1320       1330       1340       1350
GCCCTGCCCG AGTGGTCCAG TGCATTTCAG AGCTGGGCAT CTTTGGGGTC TATGTCAGGC AGGAAAAGAC ACTCGTGATG AACAAGCACG

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1360       1370       1380       1390       1400       1410       1420       1430       1440
TGGGGCATCT ACTTCGAACC AAAGCCATCG AGCATGCAGA TGGTGGTGTG GCAGCGGGAG TGGCAGTCCT GGACAACCCA TACCCTGTGT

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1450       1460       1470       1480       1490       1500       1510       1520       1530
GAGGGCACAA CCAGGCCACG GGACCTTCTA TCCTCTGTAT TTGTCATTCC TCTCCTAGCC CTCCTGAGGG GTATCCTCCT AAAGACCTCC

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1540       1550       1560       1570       1580       1590       1600       1610       1620
AAAGTTTTTA TGGAAGGGTA AATACTGGTA CCTTCCCCCA GCTTTCCATC TGAGGACCAG AAAAGTTGTG TCTCCCTTAG ATGAGATCTA

....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1630       1640       1650       1660       1670       1680       1690       1700       1710
GACGCCCCCA AATCCTTGAG ATGTGGGTAT AGCTCAGGGT AAGCTGCTCT GAGGTAAAGG TCCATGAACC CTGCCCCACT CCTGTCAGCC
```

```
....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
      1720       1730       1740       1750       1760       1770       1780       1790       1800
CCTCATCAGC CTTTTCAGCA GGTTCCAGTG CCTGACTTGG ATAGACTGAG TGGTAGGAGG AGGGGGAGTG GAGGGATAGC CTTTCCCTAA

....|....| ....|....| ....|....| ....|....| ....|....|
      1810       1820       1830       1840       1850
TTCTGCCTTA AATAAAACTG CATTGCTGAT TCAAAAAAAA AAAAAAAAAA
```

9.4 Appendix D: Sequence of GFP gene cloned into pQBI25-f

The sequence of the portion of plasmid pQBI25-f containing the GFP gene used to create GFP dsRNA for RNAi control experiments, with the primer positions highlighted.

```
 841 gtaggcgtgt acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg
 901 cctggagacg ccatccacgc tgttttgacc tccatagaag acaccgggac cgatccagcc
 961 tccgcgggcg cgcaagaaat ggctagcaaa ggagaagaac tcttcactgg agttgtccca
1021 attcttgttg aattagatgg tgatgttaac ggccacaagt tctctgtcag tggagagggt
1081 gaaggtgatg caacatacgg aaaacttacc ctgaagttca tctgcactac tggcaaactg
1141 cctgttccat ggccaacact agtcactact ctgtgctatg gtgttcaatg cttttcaaga
1201 tacccggatc atatgaaacg gcatgacttt ttcaagagtg ccatgcccga aggttatgta
1261 caggaaagga ccatcttctt caaagatgac ggcaactaca agacacgtgc tgaagtcaag
1321 tttgaaggtg atacccttgt taatagaatc gagttaaaag gtattgactt caaggaagat
1381 ggcaacattc tgggacacaa attggaatac aactataact cacacaatgt atacatcatg
1441 gcagacaaac aaaagaatgg aatcaaagtg aacttcaaga cccgccacaa cattgaagat
1501 ggaagcgttc aactagcaga ccattatcaa caaaatactc caattggcga tggccctgtc
1561 cttttaccag acaaccatta cctgtccaca caatctgccc tttcgaaaga tcccaacgaa
1621 aagagagacc acatggtcct tcttgagttt gtaacagctg ctgggattac acatggcatg
1681 gatgaactgt acaacatcga ttccactagt aacggccgcc agtgtgctgg aattctgcag
1741 atatccatca cactggcggc cgctcgagca tgcatctaga gggccctatt ctatagtgtc
1801 acctaaatgc tagagctcgc tgatcagcct cgactgtgcc ttctagttgc cagccatctg
1861 ttgtttgccc ctccccgtg ccttccttga ccctggaagg tgccactccc actgtccttt
1921 cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg
```

9.5 Appendix E: GenBank accession numbers of cDNA clones sequenced in Chapter 5

GenBank Accession number	Transcript	Source of cDNA
EF419254	T1 (CG6835-RC)	*y sn*^{3} *v* (double GS) adult fly
EF419255	T2 (CG6835-RD)	*y sn*^{3} *v* (double GS) adult fly
EF419256	T4	*y sn*^{3} *v* (double GS) adult fly
EF419257	T6 (CG32495-RA)	S2 cells
EF419258	T8 (CG32495-RB)	S2 cells
EF419259	T8 (CG32495-RB)	*w sn*^{3} (single GS) adult fly
EF419260	T8 (CG32495-RB)	*y sn*^{3} *v* (double GS) adult fly
EF419261	T10	*y sn*^{3} *v* (double GS) adult fly
EF419262	T10	*w sn*^{3} (single GS) adult fly

References

Abdollahi, M., Ranjbar, A., Shadnia, S., Nikfar, S. and Rezaiee, A. (2004). "Pesticides and oxidative stress: A review." Med Sci Monit **10**(6): RA141-7.

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X., Brandon, R. C., Rogers, Y. H., Blazej, R. G., Champe, M., Pfeiffer, B. D., Wan, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Gabor, G. L., Abril, J. F., Agbayani, A., An, H. J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E. M., Beeson, K. Y., Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K. C., Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A. D., Dew, I., Dietz, S. M., Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C., Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A. E., Garg, N. S., Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z., Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z., Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. R., Pacleb, J. M., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M. P., Smith, T., Spier, E., Spradling, A. C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wassarman, D. A., Weinstock, G. M., Weissenbach, J., Williams, S. M., WoodageT, Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H., Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O., Gibbs, R. A., Myers, E. W., Rubin, G. M. and Venter, J. C. (2000). "The genome sequence of *Drosophila melanogaster*." Science **287**(5461): 2185-95.

Adams, M. D. and Sekelsky, J. J. (2002). "From sequence to phenotype: Reverse genetics in *Drosophila melanogaster*." Nat Rev Genet **3**(3): 189-98.

Al-Lahham, A., Rohde, V., Heim, P., Leuchter, R., Veeck, J., Wunderlich, C., Wolf, K. and Zimmermann, M. (1999). "Biosynthesis of phytochelatins in the fission yeast. Phytochelatin synthesis: A second role for the glutathione synthetase gene of *Schizosaccharomyces pombe*." Yeast **15**(5): 385-96.

Allen, J. F. (1993). "Redox control of transcription: Sensors, response regulators, activators and repressors." FEBS Lett **332**(3): 203-7.

Alvarado, C., Alvarez, P., Jimenez, L. and De la Fuente, M. (2006). "Oxidative stress in leukocytes from young prematurely aging mice is reversed by supplementation with biscuits rich in antioxidants." Dev Comp Immunol **30**(12): 1168-80.

Amarzguioui, M., Rossi, J. J. and Kim, D. (2005). "Approaches for chemically synthesized siRNA and vector-mediated RNAi." FEBS Lett **579**(26): 5974-81.

Ames, B. N. (1989). "Endogenous oxidative DNA damage, aging, and cancer." Free Radic Res Commun **7**(3-6): 121-8.

Ames, B. N., Shigenaga, M. K. and Hagen, T. M. (1993). "Oxidants, antioxidants, and the degenerative diseases of aging." Proc Natl Acad Sci U S A **90**(17): 7915-22.

An, J. H. and Blackwell, T. K. (2003). "Skn-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response." Genes Dev **17**(15): 1882-93.

Anderson, M. E. (1998). "Glutathione: An overview of biosynthesis and modulation." Chem Biol Interact **111-112**: 1-14.

Arking, R., Buck, S., Berrios, A., Dwyer, S. and Baker, G. T., 3rd (1991). "Elevated paraquat resistance can be used as a bioassay for longevity in a genetically based long-lived strain of *Drosophila*." Dev Genet **12**(5): 362-70.

Arking, R., Burde, V., Graves, K., Hari, R., Feldman, E., Zeevi, A., Soliman, S., Saraiya, A., Buck, S., Vettraino, J., Sathrasala, K., Wehr, N. and Levine, R. L. (2000). "Forward and reverse selection for longevity in *Drosophila* is characterized by alteration of antioxidant gene expression and oxidative damage patterns." Exp Gerontol **35**(2): 167-85.

Armstrong, J. S. and Jones, D. P. (2002). "Glutathione depletion enforces the mitochondrial permeability transition and causes cell death in Bcl-2 overexpressing HL60 cells." Faseb J **16**(10): 1263-5.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G. and Struhl, K. (1987). Current protocols in molecular biology. New York, John Wiley & Sons.

Aw, T. Y. (2003). "Cellular redox: A modulator of intestinal epithelial cell proliferation." News Physiol Sci **18**(5): 201-204.

Back, S. A., Gan, X., Li, Y., Rosenberg, P. A. and Volpe, J. J. (1998). "Maturation-dependent vulnerability of oligodendrocytes to oxidative stress-induced death caused by glutathione depletion." J Neurosci **18**(16): 6241-53.

Bains, J. S. and Shaw, C. A. (1997). "Neurodegenerative disorders in humans: The role of glutathione in oxidative stress-mediated neuronal death." Brain Res Brain Res Rev **25**(3): 335-58.

Baldal, E. A., Brakefield, P. M. and Zwaan, B. J. (2006). "Multitrait evolution in lines of *Drosophila melanogaster* selected for increased starvation resistance: The role of metabolic rate and implications for the evolution of longevity." Evolution Int J Org Evolution **60**(7): 1435-44.

Barger, J. L., Walford, R. L. and Weindruch, R. (2003). "The retardation of aging by caloric restriction: Its significance in the transgenic era." Exp Gerontol **38**(11-12): 1343-51.

Basi, G., Schmid, E. and Maundrell, K. (1993). "Tata box mutations in the *Schizosaccharomyces pombe* nmt1 promoter affect transcription efficiency but not the transcription start point or thiamine repressibility." Gene **123**(1): 131-6.

Bauer, H., Kanzok, S. M. and Schirmer, R. H. (2002). "Thioredoxin-2 but not thioredoxin-1 is a substrate of thioredoxin peroxidase-1 from *Drosophila melanogaster*: Isolation and characterization of a second thioredoxin in *D. melanogaster* and evidence for distinct biological functions of Trx-1 and Trx-2." J Biol Chem **277**(20): 17457-63.

Bauer, J. H., Poon, P. C., Glatt-Deeley, H., Abrams, J. M. and Helfand, S. L. (2005). "Neuronal expression of p53 dominant-negative proteins in adult *Drosophila melanogaster* extends life span." Curr Biol **15**(22): 2063-8.

Baumeister, R., Schaffitzel, E. and Hertweck, M. (2006). "Endocrine signaling in *Caenorhabditis elegans* controls stress response and longevity." J Endocrinol **190**(2): 191-202.

Bayne, A. C., Mockett, R. J., Orr, W. C. and Sohal, R. S. (2005). "Enhanced catabolism of

mitochondrial superoxide/hydrogen peroxide and aging in transgenic *Drosophila*." Biochem J **391**(Pt 2): 277-84.

Becker, W., Kleinsmith, L. and Hardin, J. (2000). The world of the cell. Redwood City, California, Benjamin Cummings.

Bernards, A. and Hariharan, I. K. (2001). "Of flies and men--studying human disease in *Drosophila*." Curr Opin Genet Dev **11**(3): 274-8.

Berry, A., Capone, F., Giorgio, M., Pelicci, P. G., de Kloet, E. R., Alleva, E., Minghetti, L. and Cirulli, F. (2006). "Deletion of the life span determinant p66(shc) prevents age-dependent increases in emotionality and pain sensitivity in mice." Exp Gerontol.

Betts, M. J. and Russell, R. B. (2003). Amino acid properties and consequences of subsitutions. Bioinformatics for geneticists. M. R. Barnes and I. C. Gray. New York, Wiley

Bhole, D., Allikian, M. J. and Tower, J. (2004). "Doxycycline-regulated over-expression of hsp22 has negative effects on stress resistance and life span in adult *Drosophila melanogaster*." Mech Ageing Dev **125**(9): 651-63.

Bier, E. and Bodmer, R. (2004). "*Drosophila*, an emerging model for cardiac disease." Gene **342**(1): 1-11.

Biesalski, H. K. (2002). "Free radical theory of aging." Curr Opin Clin Nutr Metab Care **5**(1): 5-10.

Birchler, J., Pal-Bhadra, M. and Bhadra, U. (2003). Transgene cosuppression in animals. RNAi: A guide to gene silencing. G. J. Hannon. New York, Cold Spring Harbour: 23-42.

Bonilla, E., Medina-Leendertz, S. and Diaz, S. (2002). "Extension of life span and stress resistance of *Drosophila melanogaster* by long-term supplementation with melatonin." Exp Gerontol **37**(5): 629-38.

Borras, C., Esteve, J. M., Vina, J. R., Sastre, J., Vina, J. and Pallardo, F. V. (2004). "Glutathione regulates telomerase activity in 3T3 fibroblasts." J Biol Chem **279**(33): 34332-5.

Boutros, M., Kiger, A. A., Armknecht, S., Kerr, K., Hild, M., Koch, B., Haas, S. A., Consortium, H. F., Paro, R. and Perrimon, N. (2004). "Genome-wide RNAi analysis of growth and viability in *Drosophila* cells." Science **303**(5659): 832-5.

Brack, C., Bechter-Thuring, E. and Labuhn, M. (1997). "N-acetylcysteine slows down ageing and increases the life span of *Drosophila melanogaster*." Cell Mol Life Sci **53**(11-12): 960-6.

Brand, A. H. and Perrimon, N. (1993). "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes." Development **118**(2): 401-15.

Broome, C. S., Kayani, A. C., Palomero, J., Dillmann, W. H., Mestril, R., Jackson, M. J. and McArdle, A. (2006). "Effect of lifelong overexpression of hsp70 in skeletal muscle on age-related oxidative stress and adaptation after nondamaging contractile activity." Faseb J **20**(9): 1549-51.

Broughton, S. J., Piper, M. D., Ikeya, T., Bass, T. M., Jacobson, J., Driege, Y., Martinez, P., Hafen, E., Withers, D. J., Leevers, S. J. and Partridge, L. (2005). "Longer lifespan, altered metabolism, and stress resistance in *Drosophila* from ablation of cells making insulin-like ligands." Proc Natl Acad Sci U S A **102**(8): 3105-10.

Burger, J. M. and Promislow, D. E. (2004). "Sex-specific effects of interventions that extend fly life span." Sci Aging Knowledge Environ **2004**(28): pe30.

Bus, J. S. and Gibson, J. E. (1984). "Paraquat: Model for oxidant-initiated toxicity." Environ Health Perspect **55**: 37-46.

Cadenas, E. and Davies, K. J. (2000). "Mitochondrial free radical generation, oxidative stress, and aging." Free Radic Biol Med **29**(3-4): 222-30.

Caplen, N. J., Fleenor, J., Fire, A. and Morgan, R. A. (2000). "dsRNA-mediated gene silencing in cultured *Drosophila* cells: A tissue culture model for the analysis of RNA interference." Gene **252**(1-2): 95-105.

Cardozo-Pelaez, F., Stedeford, T. J., Brooks, P. J., Song, S. and Sanchez-Ramos, J. R. (2002). "Effects of diethylmaleate on DNA damage and repair in the mouse brain." Free Radic Biol Med **33**(2): 292-8.

Carey, J., Liedo, P., Oroxco, D., Tatar, M. and Vaupel, J. (1995). "A male-female longevity paradox in medfly cohorts." Journal of Animal Ecology **64**(1): 107-116.

Carmel-Harel, O. and Storz, G. (2000). "Roles of the glutathione- and thioredoxin-dependent reduction systems in the *Escherichia coli* and *Saccharomyces cerevisiae* responses to oxidative stress." Annu Rev Microbiol **54**: 439-61.

Carthew, R. W. (2003). RNAi applications in *Drosophila melanogaster*. Rnai: A guide to gene silencing. G. J. Hammond. New York, Cold Spring Harbour**:** 361-400.

Castro, J. P. and Carareto, C. M. (2004). "*Drosophila melanogaster* P transposable elements: Mechanisms of transposition and regulation." Genetica **121**(2): 107-18.

Cavener, D. R. (1987). "Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates." Nucleic Acids Res **15**(4): 1353-61.

Celniker, S. E., Wheeler, D. A., Kronmiller, B., Carlson, J. W., Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S. P., Frise, E., Hodgson, A., George, R. A., Hoskins, R. A., Laverty, T., Muzny, D. M., Nelson, C. R., Pacleb, J. M., Park, S., Pfeiffer, B. D., Richards, S., Sodergren, E. J., Svirskas, R., Tabor, P. E., Wan, K., Stapleton, M., Sutton, G. G., Venter, C., Weinstock, G., Scherer, S. E., Myers, E. W., Gibbs, R. A. and Rubin, G. M. (2002). "Finishing a whole-genome shotgun: Release 3 of the *Drosophila melanogaster* euchromatic genome sequence." Genome Biol **3**(12): RESEARCH0079.

Cho, C. G., Kim, H. J., Chung, S. W., Jung, K. J., Shim, K. H., Yu, B. P., Yodoi, J. and Chung, H. Y. (2003). "Modulation of glutathione and thioredoxin systems by calorie restriction during the aging process." Exp Gerontol **38**(5): 539-48.

Choi, J., Liu, R. M., Kundu, R. K., Sangiorgi, F., Wu, W., Maxson, R. and Forman, H. J. (2000). "Molecular mechanism of decreased glutathione content in human immunodeficiency virus type 1 tat-transgenic mice." J Biol Chem **275**(5): 3693-8.

Choi, Y. S., Lee, K. S., Yoon, H. J., Kim, I., Sohn, H. D. and Jin, B. R. (2006). "Bombus ignitus cu,zn superoxide dismutase (SOD1): cDNA cloning, gene structure, and up-regulation in response to paraquat, temperature stress, or lipopolysaccharide stimulation." Comp Biochem Physiol B Biochem Mol Biol **144**(3): 365-71.

Christiansen, L., Bathum, L., Frederiksen, H. and Christensen, K. (2004). "Paraoxonase 1 polymorphisms and survival." Eur J Hum Genet **12**(10): 843-7.

Christiansen, L., Brasch-Andersen, C., Bathum, L., Kruse, T. A. and Christensen, K. (2006). "A longitudinal study of the effect of GSST1 and GSTM1 gene copy number on survival." Mech Ageing Dev. **127**(7): 597-7.

Clancy, D. J., Gems, D., Harshman, L. G., Oldham, S., Stocker, H., Hafen, E., Leevers, S. J. and Partridge, L. (2001). "Extension of life-span by loss of chico, a *Drosophila* insulin receptor substrate protein." Science **292**(5514): 104-6.

Clancy, D. J., Gems, D., Hafen, E., Leevers, S. J. and Partridge, L. (2002). "Dietary restriction in long-lived dwarf flies." Science **296**(5566): 319.

Clemens, J. C., Worby, C. A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B. A. and Dixon, J. E. (2000). "Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways." Proc Natl Acad Sci U S A **97**(12): 6499-503.

Coe, J. P., Rahman, I., Sphyris, N., Clarke, A. R. and Harrison, D. J. (2002). "Glutathione and

p53 independently mediate responses against oxidative stress in ES cells." Free Radic Biol Med **32**(2): 187-96.
Collins, A. R. (2004). "The comet assay for DNA damage and repair: Principles, applications, and limitations." Mol Biotechnol **26**(3): 249-61.
Conant, G. C. and Wagner, A. (2003). "Asymmetric sequence divergence of duplicate genes." Genome Res **13**(9): 2052-8.
Cowan, R. (1981). The theory of atomic structure and spectra. Berkeley, University of California Press.
Cristofanon, S., Dicato, M., Ghibelli, L. and Diederich, M. (2006). "Glutathione as a mediator of apoptotic cell signaling pathways." Biochem Pharmacol.
Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W. J. and Busslinger, M. (1999). "Twin of eyeless, a second pax-6 gene of *Drosophila*, acts upstream of eyeless in the control of eye development." Mol Cell **3**(3): 297-307.
de la Asuncion, J. G., Millan, A., Pla, R., Bruseghini, L., Esteras, A., Pallardo, F. V., Sastre, J. and Vina, J. (1996). "Mitochondrial glutathione oxidation correlates with age-associated oxidative damage to mitochondrial DNA." Faseb J **10**(2): 333-8.
Dickinson, D. A., Levonen, A. L., Moellering, D. R., Arnold, E. K., Zhang, H., Darley-Usmar, V. M. and Forman, H. J. (2004). "Human glutamate cysteine ligase gene regulation through the electrophile response element." Free Radic Biol Med **37**(8): 1152-9.
Dinescu, A., Cundari, T. R., Bhansali, V. S., Luo, J. L. and Anderson, M. E. (2004). "Function of conserved residues of human glutathione synthetase: Implications for the ATP-grasp enzymes." J Biol Chem.
Dominguez, A. and Albornoz, J. (1999). "Structural instability of 297 element in *Drosophila melanogaster*." Genetica **105**(3): 239-48.
Dringen, R., Kussmaul, L., Gutterer, J. M., Hirrlinger, J. and Hamprecht, B. (1999). "The glutathione system of peroxide detoxification is less efficient in neurons than in astroglial cells." J Neurochem **72**(6): 2523-30.
Droge, W. (2002). "Free radicals in the physiological control of cell function." Physiol Rev **82**(1): 47-95.
Droge, W. (2005). "Oxidative aging and insulin receptor signaling." J Gerontol A Biol Sci Med Sci **60**(11): 1378-85.
Drysdale, R. A. and Crosby, M. A. (2005). "Flybase: Genes and gene models." Nucleic Acids Res **33**(Database issue): D390-5.
Duffy, J. B. (2002). "Gal4 system in *Drosophila*: A fly geneticist's swiss army knife." Genesis **34**(1-2): 1-15.
Dugan, L. L. and Quick, K. L. (2005). "Reactive oxygen species and aging: Evolving questions." Sci Aging Knowledge Environ **2005**(26): pe20.
Eberl, D. F., Perkins, L. A., Engelstein, M., Hilliker, A. J. and Perrimon, N. (1992). "Genetic and developmental analysis of polytene section 17 of the X chromosome of *Drosophila melanogaster*." Genetics **130**(3): 569-83.
Eklow, L., Moldeus, P. and Orrenius, S. (1984). "Oxidation of glutathione during hydroperoxide metabolism. A study using isolated hepatocytes and the glutathione reductase inhibitor 1,3-bis(2-chloroethyl)-1-nitrosourea." Eur J Biochem **138**(3): 459-63.
Engels, W. R. (1996). P-elements in *Drosophila*. Transposable elements. H. a. G. Saedler, A. Berlin, Springer-Verlag: 103-123.
Feng, J., Bussiere, F. and Hekimi, S. (2001). "Mitochondrial electron transport is a key determinant of life span in *Caenorhabditis elegans*." Dev Cell **1**(5): 633-44.
Filomeni, G., Rotilio, G. and Ciriolo, M. R. (2002). "Cell signalling and the glutathione redox system." Biochem Pharmacol **64**(5-6): 1057-64.

Finkel, T. and Holbrook, N. J. (2000). "Oxidants, oxidative stress and the biology of ageing." Nature **408**(6809): 239-47.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*." Nature **391**(6669): 806-11.

Forbes, S. N., Valenzuela, R. K., Keim, P. and Service, P. M. (2004). "Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. I. Composite interval mapping." Genetics **168**(1): 301-11.

Forman, H. J., Fukuto, J. M. and Torres, M. (2004). "Redox signaling: Thiol chemistry defines which reactive oxygen and nitrogen species can act as second messengers." Am J Physiol Cell Physiol **287**(2): C246-256.

Forsburg, S. L. (1993). "Comparison of *Schizosaccharomyces pombe* expression systems." Nucleic Acids Res **21**(12): 2955-6.

Franklin, C. C., Rosenfeld-Franklin, M. E., White, C., Kavanagh, T. J. and Fausto, N. (2003). "TGFβ1-induced suppression of glutathione antioxidant defenses in hepatocytes: Caspase-dependent post-translational and caspase-independent transcriptional regulatory mechanisms." FASEB J. **17**(11): 1535-1537.

Fraser, J. A., Saunders, R. D. and McLellan, L. I. (2002). "*Drosophila melanogaster* glutamate-cysteine ligase activity is regulated by a modifier subunit with a mechanism of action similar to that of the mammalian form." J Biol Chem **277**(2): 1158-65.

Frendo, P., Jimenez, M. J., Mathieu, C., Duret, L., Gallesi, D., Van de Sype, G., Herouart, D. and Puppo, A. (2001). "A Medicago truncatula homoglutathione synthetase is derived from glutathione synthetase by gene duplication." Plant Physiol **126**(4): 1706-15.

Fridell, Y. W., Sanchez-Blanco, A., Silvia, B. A. and Helfand, S. L. (2005). "Targeted expression of the human uncoupling protein 2 (hucp2) to adult neurons extends life span in the fly." Cell Metab **1**(2): 145-52.

Friedman, D. B. and Johnson, T. E. (1988). "A mutation in the age-1 gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility." Genetics **118**(1): 75-86.

Frosali, S., Di Simplicio, P., Perrone, S., Di Giuseppe, D., Longini, M., Tanganelli, D. and Buonocore, G. (2004). "Glutathione recycling and antioxidant enzyme activities in erythrocytes of term and preterm newborns at birth." Biology of the Neonate **85**(3): 188-194.

Gali, R. R. and Board, P. G. (1997). "Identification of an essential cysteine residue in human glutathione synthase." Biochem J **321 (Pt 1)**: 207-10.

Galperin, M. Y. and Koonin, E. V. (1997). "A diverse superfamily of enzymes with atp-dependent carboxylate-amine/thiol ligase activity." Protein Sci **6**(12): 2639-43.

Giannakou, M. E., Goss, M., Junger, M. A., Hafen, E., Leevers, S. J. and Partridge, L. (2004). "Long-lived *Drosophila* with overexpressed dFOXO in adult fat body." Science **305**(5682): 361.

Giordano, E., Rendina, R., Peluso, I. and Furia, M. (2002). "RNAi triggered by symmetrically transcribed transgenes in *Drosophila melanogaster*." Genetics **160**(2): 637-48.

Girardot, F., Monnier, V. and Tricoire, H. (2004). "Genome wide analysis of common and specific stress responses in adult *Drosophila melanogaster*." BMC Genomics **5**(1): 74.

Girardot, F., Lasbleiz, C., Monnier, V. and Tricoire, H. (2006). "Specific age related signatures in *Drosophila* body parts transcriptome." BMC Genomics **7**(1): 69.Godoy-Herrera, R., Burnet, B. and Connolly, K. (2004). "Conservation and divergence of the genetic structure of larval foraging behaviour in two species of the *Drosophila* simulans clade." Heredity **92**(1): 14-9.

Grant, C. M., MacIver, F. H. and Dawes, I. W. (1997). "Glutathione synthetase is dispensable for growth under both normal and oxidative stress conditions in the yeast

Saccharomyces cerevisiae due to an accumulation of the dipeptide gamma-glutamylcysteine." Mol Biol Cell **8**(9): 1699-707.
Gredilla, R. and Barja, G. (2005). "Minireview: The role of oxidative stress in relation to caloric restriction and longevity." Endocrinology **146**(9): 3713-7.
Grumbling, G. and Strelets, V. (2006). "Flybase: Anatomical data, images and queries." Nucleic Acids Res **34**(Database issue): D484-8.
Guarente, L. and Kenyon, C. (2000). "Genetic pathways that regulate ageing in model organisms." Nature **408**(6809): 255-62.
Guarente, L. and Picard, F. (2005). "Calorie restriction--the SIR2 connection." Cell **120**(4): 473-82.
Gushima, H., Miya, T., Murata, K. and Kimura, A. (1983). "Construction of glutathione-producing strains of *Escherichia coli* b by recombinant DNA techniques." J Appl Biochem **5**(1-2): 43-52.
Gushima, H., Yasuda, S., Soeda, E., Yokota, M., Kondo, M. and Kimura, A. (1984). "Complete nucleotide sequence of the E. coli glutathione synthetase GSH-II." Nucleic Acids Res **12**(24): 9299-307.
Ha, E. M., Oh, C. T., Ryu, J. H., Bae, Y. S., Kang, S. W., Jang, I. H., Brey, P. T. and Lee, W. J. (2005). "An antioxidant system required for host protection against gut infection in *Drosophila*." Dev Cell **8**(1): 125-32.
Hall, T. A. (1999). " BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt." Nucl. Acids. Symp. Ser **41**: 95-98.
Halliwell, B. and Gutteridge, J. M. C. (1999). Free radicals in biology and medicine. New York, Oxford science publications.
Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). "An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells." Nature **404**(6775): 293-6.
Hammond, S. M. (2003). RNAi technologies in *Drosophila* cell culture. RNAi: A guide to gene silencing. G. Hammond. New York, Cold Spring Harbour: 345-360.
Hammond, S. M. (2005). "Dicing and slicing: The core machinery of the RNA interference pathway." FEBS Lett **579**(26): 5822-9.
Hancock, J. T., Desikan, R. and Neill, S. J. (2001). "Does the redox status of cytochrome C act as a fail-safe mechanism in the regulation of programmed cell death?" Free Radic Biol Med **31**(5): 697-703.
Hannon, G. J. (2002). "RNA interference." Nature **418**(6894): 244-51.
Harman, D. (1956). "Aging: A theory based on free radical and radiation chemistry." J Gerontol **11**(3): 298-300.
Harman, D. (1972). "The biologic clock: The mitochondria?" J Am Geriatr Soc **20**(4): 145-7.
Harman, D. (1981). "The aging process." Proc Natl Acad Sci U S A **78**(11): 7124-8.
Harman, D. (2006). "Free radical theory of aging: An update: Increasing the functional life span." Ann N Y Acad Sci **1067**: 10-21.
Hasty, P., Campisi, J., Hoeijmakers, J., van Steeg, H. and Vijg, J. (2003). "Aging and genome maintenance: Lessons from the mouse?" Science **299**(5611): 1355-9.
Hazzard, W. R. (1986). "Biological basis of the sex differential in longevity." J Am Geriatr Soc **34**(6): 455-71.
Hekimi, S. and Guarente, L. (2003). "Genetics and the specificity of the aging process." Science **299**(5611): 1351-4.
Helfand, S. L. and Rogina, B. (2003a). "From genes to aging in *Drosophila*." Adv Genet **49**: 67-109.
Helfand, S. L. and Rogina, B. (2003b). "Molecular genetics of aging in the fly: Is this the end of the beginning?" Bioessays **25**(2): 134-41.

Herndon, L. A., Schmeissner, P. J., Dudaronek, J. M., Brown, P. A., Listner, K. M., Sakano, Y., Paupard, M. C., Hall, D. H. and Driscoll, M. (2002). "Stochastic and genetic factors influence tissue-specific decline in ageing *C. elegans*." Nature **419**(6909): 808-14.

Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P. C., Cervera, P. and Le Bouc, Y. (2003). "IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice." Nature **421**(6919): 182-7.

Holzenberger, M., Kappeler, L. and De Magalhaes Filho, C. (2004). "IGF-1 signaling and aging." Exp Gerontol **39**(11-12): 1761-4.

Hong, H., Lu, Y., Ji, Z.-N. and Liu, G.-Q. (2006). "Up-regulation of p-glycoprotein expression by glutathione depletion-induced oxidative stress in rat brain microvessel endothelial cells." Journal of Neurochemistry **98**(5): 1465-1473.

Hosono, R., Nishimoto, S. and Kuno, S. (1989). "Alterations of life span in the nematode *Caenorhabditis elegans* under monoxenic culture conditions." Exp Gerontol **24**(3): 251-64.

Huang, C. S., Anderson, M. E. and Meister, A. (1993a). "Amino acid sequence and function of the light subunit of rat kidney gamma-glutamylcysteine synthetase." J. Biol. Chem. **268**(27): 20578-20583.

Huang, C. S., Chang, L. S., Anderson, M. E. and Meister, A. (1993b). "Catalytic and regulatory properties of the heavy subunit of rat kidney gamma-glutamylcysteine synthetase." J. Biol. Chem. **268**(26): 19675-19680.

Huang, C. S., He, W., Meister, A. and Anderson, M. E. (1995). "Amino acid sequence of rat kidney glutathione synthetase." Proc Natl Acad Sci U S A **92**(4): 1232-6.

Huang, Z. A., Yang, H., Chen, C., Zeng, Z. and Lu, S. C. (2000). "Inducers of gamma-glutamylcysteine synthetase and their effects on glutathione synthetase expression." Biochim Biophys Acta **1493**(1-2): 48-55.

Hulbert, A. J., Clancy, D. J., Mair, W., Braeckman, B. P., Gems, D. and Partridge, L. (2004). "Metabolic rate is not reduced by dietary-restriction or by lowered insulin/IGF-1 signalling and is not correlated with individual lifespan in *Drosophila melanogaster*." Exp Gerontol **39**(8): 1137-43.

Hurles, M. (2004). "Gene duplication: The genomic trade in spare parts." PLoS Biol **2**(7): E206.

Hwangbo, D. S., Gersham, B., Tu, M. P., Palmer, M. and Tatar, M. (2004). "*Drosophila* dFOXO controls lifespan and regulates insulin signalling in brain and fat body." Nature **429**(6991): 562-6.

Iles, K. E. and Liu, R. M. (2005). "Mechanisms of glutamate cysteine ligase (GCL) induction by 4-hydroxynonenal." Free Radic Biol Med **38**(5): 547-56.

Inoue, Y., Sugiyama, K., Ueminami, H., Izawa, S. and Kimura, A. (1996). "Modification of *Escherichia coli* b glutathione synthetase with polyethylene glycol for clinical application to enzyme replacement therapy for glutathione deficiency." Clin Diagn Lab Immunol **3**(6): 663-8.

Inoue, Y., Sugiyama, K., Izawa, S. and Kimura, A. (1998). "Molecular identification of glutathione synthetase (gsh2) gene from *Saccharomyces cerevisiae*." Biochim Biophys Acta **1395**(3): 315-20.

Ito, H., Okamoto, K. and Kato, K. (1998). "Enhancement of expression of stress proteins by agents that lower the levels of glutathione in cells." Biochim Biophys Acta **1397**(2): 223-30.

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D. and Yamamoto, M. (1999). "Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal neh2 domain." Genes Dev **13**(1): 76-86.

Jackson, A. U., Galecki, A. T., Burke, D. T. and Miller, R. A. (2002). "Mouse loci associated with life span exhibit sex-specific and epistatic effects." J Gerontol A Biol Sci Med Sci **57**(1): B9-B15.

Janaky, R., Ogita, K., Pasqualotto, B. A., Bains, J. S., Oja, S. S., Yoneda, Y. and Shaw, C. A. (1999). "Glutathione and signal transduction in the mammalian CNS." J Neurochem **73**(3): 889-902.

Janowiak, B. E. and Griffith, O. W. (2005). "Glutathione synthesis in Streptococcus agalactiae. One protein accounts for gamma-glutamylcysteine synthetase and glutathione synthetase activities." J Biol Chem **280**(12): 11829-39.

Jazwinski, S. M. and Conzelmann, A. (2002). "LAG1 puts the focus on ceramide signaling." Int J Biochem Cell Biol **34**(11): 1491-5.

Jez, J. M. and Cahoon, R. E. (2004). "Kinetic mechanism of glutathione synthetase from *Arabidopsis thaliana*." J Biol Chem **279**(41): 42726-31.

Johnson, T. K., McKechnie, S. W. and Clancy, D. J. (2006). "Water balance and cation levels in *Drosophila*: Can early physiological decline predict aging and longevity?" J Gerontol A Biol Sci Med Sci **61**(2): 146-52.

Jones, D. P. (2002). "Redox potential of GSH/GSSG couple: Assay and biological significance." Methods Enzymol **348**: 93-112.

Kang, H. L., Benzer, S. and Min, K. T. (2002). "Life extension in *Drosophila* by feeding a drug." Proc Natl Acad Sci U S A **99**(2): 838-43.

Kanzok, S. M., Fechner, A., Bauer, H., Ulschmid, J. K., Muller, H. M., Botella-Munoz, J., Schneuwly, S., Schirmer, R. and Becker, K. (2001). "Substitution of the thioredoxin system for glutathione reductase in *Drosophila melanogaster*." Science **291**(5504): 643-6.

Kavi, H. H., Fernandez, H. R., Xie, W. and Birchler, J. A. (2005). "RNA silencing in *Drosophila*." FEBS Lett **579**(26): 5940-9.

Keaney, M., Matthijssens, F., Sharpe, M., Vanfleteren, J. and Gems, D. (2004). "Superoxide dismutase mimetics elevate superoxide dismutase activity in vivo but do not retard aging in the nematode *Caenorhabditis elegans*." Free Radical Biology and Medicine **37**(2): 239-250.

Kehrer, J. P. and Lund, L. G. (1994). "Cellular reducing equivalents and oxidative stress." Free Radic Biol Med **17**(1): 65-75.

Kenyon, C., Chang, J., Gensch, E., Rudner, A. and Tabtiang, R. (1993). "A *C. elegans* mutant that lives twice as long as wild type." Nature **366**(6454): 461-4.

Kim, H. G., Hong, S. M., Kim, S. J., Park, H. J., Jung, H. I., Lee, Y. Y., Moon, J. S., Lim, H. W., Park, E. H. and Lim, C. J. (2003). "Age-related changes in the activity of antioxidant and redox enzymes in rats." Mol Cells **16**(3): 278-84.

Kim, S. J., Shin, Y. H., Kim, K., Park, E. H., Sa, J. H. and Lim, C. J. (2003). "Regulation of the gene encoding glutathione synthetase from the fission yeast." J Biochem Mol Biol **36**(3): 326-31.

Kim, S. J., Kim, H. G., Kim, B. C., Park, E. H. and Lim, C. J. (2004). "Transcriptional regulation of glutathione synthetase in the fission yeast *Schizosaccharomyces pombe*." Mol Cells **18**(2): 242-8.

Kim, Y. (2004). "Collaborative effects of Photobacterium CuZn superoxide dismutase : SODs and human AP endonuclease in DNA repair and SOD-deficient *Escherichia coli* under oxidative stress." Free Radic Biol Med **36**(2): 173-9.

Kirkwood, T. B. (1977). "Evolution of ageing." Nature **270**(5635): 301-4.

Kirkwood, T. B. and Austad, S. N. (2000). "Why do we age?" Nature **408**(6809): 233-8.

Kirkwood, T. B. and Finch, C. E. (2002). "Ageing: The old worm turns more slowly." Nature **419**(6909): 794-5.

Kistler, M., Summer, K. H. and Eckardt, F. (1986). "Isolation of glutathione-deficient mutants of the yeast *Saccharomyces cerevisiae*." Mutat Res **173**(2): 117-20.
Knight, J. A. (2000). "The biochemistry of aging." Adv Clin Chem **35**: 1-62.
Kobayashi, M., Itoh, K., Suzuki, T., Osanai, H., Nishikawa, K., Katoh, Y., Takagi, Y. and Yamamoto, M. (2002). "Identification of the interactive interface and phylogenic conservation of the Nrf2-Keap1 system." Genes Cells **7**(8): 807-20.
Komlosh, A., Volohonsky, G., Porat, N., Tuby, C., Bluvshtein, E., Steinberg, P., Oesch, F. and Stark, A. A. (2002). "γ-glutamyl transpeptidase and glutathione biosynthesis in non-tumorigenic and tumorigenic rat liver oval cell lines." Carcinogenesis **23**(4): 669-678.
Kondo, T., Yoshida, K., Urata, Y., Goto, S., Gasa, S. and Taniguchi, N. (1993). "Gamma-glutamylcysteine synthetase and active transport of glutathione s-conjugate are responsive to heat shock in k562 erythroid cells." J Biol Chem **268**(27): 20366-72.
Kraut, R., Menon, K. and Zinn, K. (2001). "A gain-of-function screen for genes controlling motor axon guidance and synaptogenesis in *Drosophila*." Curr Biol **11**(6): 417-30.
Kubo, C., Day, N. K. and Good, R. A. (1984). "Influence of early or late dietary restriction on life span and immunological parameters in MRl/MR-lPR/lPR mice." Proc Natl Acad Sci U S A **81**(18): 5831-5.
Landis, G. N., Bhole, D. and Tower, J. (2003). "A search for doxycycline-dependent mutations that increase *Drosophila melanogaster* life span identifies the sugar baby, filamin, Fwd and CCTL genes." Genome Biol **4**(2): R8.
Landis, G. N., Abdueva, D., Skvortsov, D., Yang, J., Rabin, B. E., Carrick, J., Tavare, S. and Tower, J. (2004). "Similar gene expression patterns characterize aging and oxidative stress in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **101**(20): 7663-8.
Landis, G. N. and Tower, J. (2005). "Superoxide dismutase evolution and life span regulation." Mech Ageing Dev **126**(3): 365-79.
Lang, C. A., Mills, B. J., Lang, H. L., Liu, M. C., Usui, W. M., Richie, J., Jr., Mastropaolo, W. and Murrell, S. A. (2002). "High blood glutathione levels accompany excellent physical and mental health in women ages 60 to 103 years." J Lab Clin Med **140**(6): 413-7.
Le Bourg, E. (2001). "Oxidative stress, aging and longevity in *Drosophila melanogaster*." FEBS Lett **498**(2-3): 183-6.
Lee, J. H., Lee, E., Park, J., Kim, E., Kim, J. and Chung, J. (2003). "In vivo p53 function is indispensable for DNA damage-induced apoptotic signaling in *Drosophila*." FEBS Lett **550**(1-3): 5-10.
Lee, Y. S. and Carthew, R. W. (2003). "Making a better RNAi vector for *Drosophila*: Use of intron spacers." Methods **30**(4): 322-9.
Lenton, K. J., Therriault, H., Cantin, A. M., Fulop, T., Payette, H. and Wagner, J. R. (2000). "Direct correlation of glutathione and ascorbate and their dependence on age and season in human lymphocytes." Am J Clin Nutr **71**(5): 1194-200.
Lesser, K. J., Paiusi, I. C. and Leips, J. (2006). "Naturally occurring genetic variation in the age-specific immune response of *Drosophila melanogaster*." Aging Cell.
Liang, H., Masoro, E. J., Nelson, J. F., Strong, R., McMahan, C. A. and Richardson, A. (2003). "Genetic mouse models of extended lifespan." Exp Gerontol **38**(11-12): 1353-64.
Liao, G.-c., Rehm, E. J. and Rubin, G. M. (2000). "Insertion site preferences of the P transposable element in *Drosophila melanogaster*." PNAS **97**(7): 3347-3351.
Liao, X. Y., Shen, W., Chen, J., Li, Y. and Du, G. C. (2006). "Improved glutathione production by gene expression in *Escherichia coli*." Lett Appl Microbiol **43**(2): 211-4.
Lin, S. J., Defossez, P. A. and Guarente, L. (2000). "Requirement of NAD and Sir2 for life-span extension by calorie restriction in *Saccharomyces cerevisiae*." Science **289**(5487):

2126-8.
Lin, Y. J., Seroude, L. and Benzer, S. (1998). "Extended life-span and stress resistance in the *Drosophila* mutant methuselah." Science **282**(5390): 943-6.
Lindsley, D. L. and Zimm, G. G. (1992). The genome of *Drosophila melanogaster*. London, Academic Press, Inc.
Linnane, A. W. and Eastwood, H. (2006). "Cellular redox regulation and prooxidant signaling systems: A new perspective on the free radical theory of aging." Ann N Y Acad Sci **1067**: 47-55.
Lints, F. A., Bourgois, M., Delalieux, A., Stoll, J. and Lints, C. V. (1983). "Does the female life span exceed that of the male? A study in *Drosophila melanogaster*." Gerontology **29**(5): 336-52.
Lints, F. A. (1989). "The rate of living theory revisited." Gerontology **35**(1): 36-57.
Liu, H., Wang, H., Shenvi, S., Hagen, T. M. and Liu, R. M. (2004). "Glutathione metabolism during aging and in Alzheimer disease." Ann N Y Acad Sci **1019**: 346-9.
Liu, L., Wylie, R. C., Andrews, L. G. and Tollefsbol, T. O. (2003). "Aging, cancer and nutrition: The DNA methylation connection." Mech Ageing Dev **124**(10-12): 989-98.
Lomaestro, B. M. and Malone, M. (1995). "Glutathione in health and disease: Pharmacotherapeutic issues." Ann Pharmacother **29**(12): 1263-73.
Lum, L., Yao, S., Mozer, B., Rovescalli, A., Von Kessler, D., Nirenberg, M. and Beachy, P. A. (2003). "Identification of hedgehog pathway components by RNAi in *Drosophila* cultured cells." Science **299**(5615): 2039-45.
Luo, J. L., Hammarqvist, F., Andersson, K. and Wernerman, J. (1998). "Surgical trauma decreases glutathione synthetic capacity in human skeletal muscle tissue." Am J Physiol **275**(2 Pt 1): E359-65.
Lynch, M. and Conery, J. S. (2000). "The evolutionary fate and consequences of duplicate genes." Science **290**(5494): 1151-5.
Lynch, M. and Force, A. (2000). "The probability of duplicate gene preservation by subfunctionalization." Genetics **154**(1): 459-73.
Lynch, M. (2002). "Genomics. Gene duplication and evolution." Science **297**(5583): 945-7.
Mackay, W. J. and Bewley, G. C. (1989). "The genetics of catalase in *Drosophila melanogaster*: Isolation and characterization of acatalasemic mutants." Genetics **122**(3): 643-52.
Magwere, T., Chapman, T. and Partridge, L. (2004). "Sex differences in the effect of dietary restriction on life span and mortality rates in female and male *Drosophila melanogaster*." J Gerontol A Biol Sci Med Sci **59**(1): 3-9.
Magwere, T., Pamplona, R., Miwa, S., Martinez-Diaz, P., Portero-Otin, M., Brand, M. D. and Partridge, L. (2006a). "Flight activity, mortality rates, and lipoxidative damage in *Drosophila*." J Gerontol A Biol Sci Med Sci **61**(2): 136-45.
Magwere, T., West, M., Riyahi, K., Murphy, M. P., Smith, R. A. and Partridge, L. (2006b). "The effects of exogenous antioxidants on lifespan and oxidative stress resistance in *Drosophila melanogaster*." Mech Ageing Dev **127**(4): 356-70.
Mair, W., Goymer, P., Pletcher, S. D. and Partridge, L. (2003). "Demography of dietary restriction and death in *Drosophila*." Science **301**(5640): 1731-3.
Mair, W., Piper, M. D. and Partridge, L. (2005). "Calories do not explain extension of life span by dietary restriction in *Drosophila*." PLoS Biol **3**(7): e223.
Marchler-Bauer, A. and Bryant, S. H. (2004). "Cd-search: Protein domain annotations on the fly." Nucleic Acids Res **32**(Web Server issue): W327-31.
Masoro, E. (1993). Theories of ageing. Free radicals in ageing. B. Yu. Boca Raton, CRC Press.
Masoro, E. J. (2005). "Overview of caloric restriction and ageing." Mech Ageing Dev **126**(9):

913-22.
Matamoros, M. A., Moran, J. F., Iturbe-Ormaetxe, I., Rubio, M. C. and Becana, M. (1999). "Glutathione and homoglutathione synthesis in legume root nodules." Plant Physiol. **121**(3): 879-888.
Mathers, J., Fraser, J. A., McMahon, M., Saunders, R. D., Hayes, J. D. and McLellan, L. I. (2004). "Antioxidant and cytoprotective responses to redox stress." Biochem Soc Symp(71): 157-76.
Mattison, J. A., Lane, M. A., Roth, G. S. and Ingram, D. K. (2003). "Calorie restriction in Rhesus monkeys." Exp Gerontol **38**(1-2): 35-46.
Maynard Smith, J. (1962). "Review lectures on senescence. I. The causes of ageing." Proceedings of the Royal Society of London. Series B, Biological Sciences **157**(966): 115-127.
McCarroll, S. A., Murphy, C. T., Zou, S., Pletcher, S. D., Chin, C. S., Jan, Y. N., Kenyon, C., Bargmann, C. I. and Li, H. (2004). "Comparing genomic expression patterns across species identifies shared transcriptional profile in aging." Nat Genet **36**(2): 197-204.
McElwee, J. J., Schuster, E., Blanc, E., Thomas, J. H. and Gems, D. (2004). "Shared transcriptional signature in *Caenorhabditis elegans* dauer larvae and long-lived daf-2 mutants implicates detoxification system in longevity assurance." J Biol Chem **279**(43): 44533-43.
McManus, M. T. and Sharp, P. A. (2002). "Gene silencing in mammals by small interfering RNAs." Nat Rev Genet **3**(10): 737-47.
Meierjohann, S., Walter, R. D. and Muller, S. (2002). "Glutathione synthetase from *Plasmodium falciparum*." Biochem J **363**(Pt 3): 833-8.
Meister, A. and Anderson, M. E. (1983). "Glutathione." Annu Rev Biochem **52**: 711-60.
Meister, A. (1985). "Glutathione synthetase from rat kidney." Methods Enzymol **113**: 393-9.
Meister, A. (1994). "Glutathione-ascorbic acid antioxidant system in animals." J Biol Chem **269**(13): 9397-400.
Meister, A. (1995). "Mitochondrial changes associated with glutathione deficiency." Biochim Biophys Acta **1271**(1): 35-42.
Meister, A. and Larsson, A. (1995). Glutathione synthetase deficiency and other disorders of the gamma-glutamyl cycle. The metabolic and molecular bases of inherited disease. C.E.A. Scriver. New York, McGraw-Hill. **1:** 1461-1495 (chapter 43).
Melov, S., Ravenscroft, J., Malik, S., Gill, M. S., Walker, D. W., Clayton, P. E., Wallace, D. C., Malfroy, B., Doctrow, S. R. and Lithgow, G. J. (2000). "Extension of life-span with superoxide dismutase/catalase mimetics." Science **289**(5484): 1567-9.
Melov, S. and Hubbard, A. (2004). "Microarrays as a tool to investigate the biology of aging: A retrospective and a look to the future." Sci Aging Knowledge Environ **2004**(42): re7.
Melvin, R. G. and Ballard, J. W. (2006). "Intraspecific variation in survival and mitochondrial oxidative phosphorylation in wild-caught *Drosophila simulans*." Aging Cell **5**(3): 225-33.
Merry, B. J. (2004). "Oxidative stress and mitochondrial function with aging--the effects of calorie restriction." Aging Cell **3**(1): 7-12.
Miah, A., Elaswarapu, R., Cooper, M., Bell, G., Sinka, R. and Glover, D. (2004). "Use of RNAi to examine gene function in *Drosophila melanogaster*." American Biotechnology Laboratory: 12-14.
Min, K. J. and Tatar, M. (2006). "Restriction of amino acids extends lifespan in *Drosophila melanogaster*." Mech Ageing Dev **127**(7): 643-6.
Misra, S., Crosby, M. A., Mungall, C. J., Matthews, B. B., Campbell, K. S., Hradecky, P., Huang, Y., Kaminker, J. S., Millburn, G. H., Prochnik, S. E., Smith, C. D., Tupy, J. L., Whitfied, E. J., Bayraktaroglu, L., Berman, B. P., Bettencourt, B. R., Celniker, S. E.,

de Grey, A. D., Drysdale, R. A., Harris, N. L., Richter, J., Russo, S., Schroeder, A. J., Shu, S. Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W. M., Rubin, G. M. and Lewis, S. E. (2002). "Annotation of the *Drosophila melanogaster* euchromatic genome: A systematic review." Genome Biol **3**(12): RESEARCH0083.

Missirlis, F., Phillips, J. P. and Jackle, H. (2001). "Cooperative action of antioxidant defense systems in *Drosophila*." Curr Biol **11**(16): 1272-7.

Mockett, R. J., Orr, W. C., Rahmandar, J. J., Benes, J. J., Radyuk, S. N., Klichko, V. I. and Sohal, R. S. (1999a). "Overexpression of Mn-containing superoxide dismutase in transgenic *Drosophila melanogaster*." Arch Biochem Biophys **371**(2): 260-9.

Mockett, R. J., Sohal, R. S. and Orr, W. C. (1999b). "Overexpression of glutathione reductase extends survival in transgenic *Drosophila melanogaster* under hyperoxia but not normoxia." Faseb J **13**(13): 1733-42.

Mockett, R. J., Bayne, A. C., Kwong, L. K., Orr, W. C. and Sohal, R. S. (2003). "Ectopic expression of catalase in *Drosophila* mitochondria increases stress resistance but not longevity." Free Radic Biol Med **34**(2): 207-17.

Mockett, R. J. and Sohal, R. S. (2006). "Temperature-dependent trade-offs between longevity and fertility in the *Drosophila* mutant, methuselah." Exp Gerontol **41**(6): 566-73.

Moller, P. and Loft, S. (2006). "Dietary antioxidants and beneficial effect on oxidatively damaged DNA." Free Radic Biol Med **41**(3): 388-415.

Monje-Casas, F., Michan, C. and Pueyo, C. (2004). "Absolute transcript levels of thioredoxin- and glutathione-dependent redox systems in *Saccharomyces cerevisiae*: Response to stress and modulation with growth." Biochem J **383**(Pt 1): 139-47.

Morey, M., Serras, F. and Corominas, M. (2003). "Halving the selenophosphate synthetase gene dose confers hypersensitivity to oxidative stress in *Drosophila melanogaster*." FEBS Lett **534**(1-3): 111-4.

Morozova, N., Forry, E. P., Shahid, E., Zavacki, A. M., Harney, J. W., Kraytsberg, Y. and Berry, M. J. (2003). "Antioxidant function of a novel selenoprotein in *Drosophila melanogaster*." Genes Cells **8**(12): 963-71.

Morrow, G., Samson, M., Michaud, S. and Tanguay, R. M. (2004). "Overexpression of the small mitochondrial hsp22 extends *Drosophila* life span and increases resistance to oxidative stress." Faseb J **18**(3): 598-9.

Murzin, A. G. (1996). "Structural classification of proteins: New superfamilies." Curr Opin Struct Biol **6**(3): 386-94.

Mutoh, N. and Hayashi, Y. (1988). "Isolation of mutants of *Schizosaccharomyces pombe* unable to synthesize cadystin, small cadmium-binding peptides." Biochem Biophys Res Commun **151**(1): 32-9.

Mutoh, N., Nakagawa, C. W., Ando, S., Tanabe, K. and Hayashi, Y. (1991). "Cloning and sequencing of the gene encoding the large subunit of glutathione synthetase of *Schizosaccharomyces pombe*." Biochem Biophys Res Commun **181**(1): 430-6.

Nguyen, T., Sherratt, P. J. and Pickett, C. B. (2003). "Regulatory mechanisms controlling gene expression mediated by the antioxidant response element." Annual Review of Pharmacology and Toxicology **43**(1): 233-260.

Njalsson, R. (2005). "Glutathione synthetase deficiency." Cell Mol Life Sci **62**(17): 1938-45.

Njalsson, R. and Norgren, S. (2005). "Physiological and pathological aspects of GSH metabolism." Acta Paediatr **94**(2): 132-7.

Nygren, J., Ristoff, E., Carlsson, K., Moller, L. and Larsson, A. (2005). "Oxidative DNA damage in cultured fibroblasts from patients with hereditary glutathione synthetase deficiency." Free Radic Res **39**(6): 595-601.

Olive, P. L. and Banath, J. P. (1993). "Detection of DNA double-strand breaks through the cell cycle after exposure to X-rays, bleomycin, etoposide and 125IdURD." Int J Radiat Biol

64(4): 349-58.
Orr, W. C. and Sohal, R. S. (1992). "The effects of catalase gene overexpression on life span and resistance to oxidative stress in transgenic *Drosophila melanogaster*." Arch Biochem Biophys **297**(1): 35-41.
Orr, W. C. and Sohal, R. S. (1993). "Effects of Cu-Zn superoxide dismutase overexpression of life span and resistance to oxidative stress in transgenic *Drosophila melanogaster*." Arch Biochem Biophys **301**(1): 34-40.
Orr, W. C. and Sohal, R. S. (1994). "Extension of life-span by overexpression of superoxide dismutase and catalase in *Drosophila melanogaster*." Science **263**(5150): 1128-30.
Orr, W. C., Orr, E. C., Legan, S. K. and Sohal, R. S. (1996). "Molecular analysis of the *Drosophila* catalase gene." Arch Biochem Biophys **330**(2): 251-8.
Orr, W. C., Mockett, R. J., Benes, J. J. and Sohal, R. S. (2003). "Effects of overexpression of copper-zinc and manganese superoxide dismutases, catalase, and thioredoxin reductase genes on longevity in *Drosophila melanogaster*." J Biol Chem **278**(29): 26418-22.
Orr, W. C. and Sohal, R. S. (2003). "Does overexpression of Cu,Zn-SOD extend life span in *Drosophila melanogaster*?" Exp Gerontol **38**(3): 227-30.
Orr, W. C., Radyuk, S. N., Prabhudesai, L., Toroser, D., Benes, J. J., Luchak, J. M., Mockett, R. J., Rebrin, I., Hubbard, J. G. and Sohal, R. S. (2005). "Overexpression of glutamate-cysteine ligase extends life span in *Drosophila melanogaster*." J Biol Chem **280**(45): 37331-8.
Osterwalder, T., Yoon, K. S., White, B. H. and Keshishian, H. (2001). "A conditional tissue-specific transgene expression system using inducible Gal4." Proc Natl Acad Sci U S A **98**(22): 12596-601.
Pallant, J. (2005). SPSS survival manual. Maidenhead, Open University Press.
Palmeira, C. M. (1999). "Herbicide-induced mitochondrial and cellular liver toxicity: A review of paraquat, dinoseb, and 2,4-d effects." Toxic Substance Mechanisms **18**(4): 187-204.
Park, S. K. and Prolla, T. A. (2005). "Lessons learned from gene expression profile studies of aging and caloric restriction." Ageing Res Rev **4**(1): 55-65.
Parkes, T. L., Elia, A. J., Dickinson, D., Hilliker, A. J., Phillips, J. P. and Boulianne, G. L. (1998). "Extension of *Drosophila* lifespan by overexpression of human SOD1 in motorneurons." Nat Genet **19**(2): 171-4.
Parkinson, A. (2001). Biotransformation of xenobiotics. Casarett and Doull's toxicology: The basic science of poisons. C. D. Klaassen. New York, McGraw-Hill: 133-225.
Pearl, R. (1928). The rate of living, being an account of some experimental studies on the biology of life duration. New York, Alfred A. Knopf.
Pennisi, E. (1998). "Single gene controls fruit fly life-span." Science **282**(5390): 856.
Phillips, J. P., Campbell, S. D., Michaud, D., Charbonneau, M. and Hilliker, A. J. (1989). "Null mutation of copper/zinc superoxide dismutase in *Drosophila* confers hypersensitivity to paraquat and reduced longevity." Proc Natl Acad Sci U S A **86**(8): 2761-5.
Phillips, J. P., Parkes, T. L. and Hilliker, A. J. (2000). "Targeted neuronal gene expression and longevity in *Drosophila*." Exp Gerontol **35**(9-10): 1157-64.
Phlippen, N., Hoffmann, K., Fischer, R., Wolf, K. and Zimmermann, M. (2003). "The glutathione synthetase of *Schizosaccharomyces pombe* is synthesized as a homodimer but retains full activity when present as a heterotetramer." J Biol Chem **278**(41): 40152-61.
Pletcher, S. D., Macdonald, S. J., Marguerie, R., Certa, U., Stearns, S. C., Goldstein, D. B. and Partridge, L. (2002). "Genome-wide transcript profiles in aging and calorically restricted *Drosophila melanogaster*." Curr Biol **12**(9): 712-23.

Poirier, L. and Seroude, L. (2005). "Genetic approaches to study aging in *Drosophila melanogaster*." AGE **27**: 165-182.

Polekhina, G., Board, P. G., Gali, R. R., Rossjohn, J. and Parker, M. W. (1999). "Molecular basis of glutathione synthetase deficiency and a rare gene permutation event." Embo J **18**(12): 3204-13.

Prado, A., Canal, I. and Ferrus, A. (1999). "The haplolethal region at the 16F gene cluster of *Drosophila melanogaster*: Structure and function." Genetics **151**(1): 163-75.

Quarrie, J. K. and Riabowol, K. T. (2004). "Murine models of life span extension." Sci Aging Knowledge Environ **2004**(31): re5.

Radyuk, S. N., Klichko, V. I., Spinola, B., Sohal, R. S. and Orr, W. C. (2001). "The peroxiredoxin gene family in *Drosophila melanogaster*." Free Radic Biol Med **31**(9): 1090-100.

Radyuk, S. N., Sohal, R. S. and Orr, W. C. (2003). "Thioredoxin peroxidases can foster cytoprotection or cell death in response to different stressors: Over- and under-expression of thioredoxin peroxidase in *Drosophila* cells." Biochem J **371**(Pt 3): 743-52.

Radyuk, S. N., Michalak, K., Rebrin, I., Sohal, R. S. and Orr, W. C. (2006). "Effects of ectopic expression of *Drosophila* DNA glycosylases Dogg1 and Rps3 in mitochondria." Free Radical Biology and Medicine **41**(5): 757-764.

Rahman, I., Biswas, S. K., Jimenez, L. A., Torres, M. and Forman, H. J. (2005). "Glutathione, stress responses, and redox signaling in lung inflammation." Antioxid Redox Signal **7**(1-2): 42-59.

Rathbun, W. B. and Holleschau, A. M. (1992). "The effects of age on glutathione synthesis enzymes in lenses of old world simians and prosimians." Curr Eye Res **11**(7): 601-7.

Rawlins, M. R., Leaver, C. J. and May, M. J. (1995). "Characterisation of an *Arabidopsis thaliana* cDNA encoding glutathione synthetase." FEBS Lett **376**(1-2): 81-6.

Rebrin, I., Kamzalov, S. and Sohal, R. S. (2003). "Effects of age and caloric restriction on glutathione redox state in mice." Free Radic Biol Med **35**(6): 626-35.

Rebrin, I., Bayne, A. C., Mockett, R. J., Orr, W. C. and Sohal, R. S. (2004). "Free aminothiols, glutathione redox state and protein mixed disulphides in aging *Drosophila melanogaster*." Biochem J **382**(Pt 1): 131-6.

Reiter, L. T., Potocki, L., Chien, S., Gribskov, M. and Bier, E. (2001). "A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*." Genome Res **11**(6): 1114-25.

Reliene, R. and Schiestl, R. H. (2005). "Glutathione depletion by buthionine sulfoximine induces DNA deletions in mice." Carcinogenesis.

Reveillaud, I., Niedzwiecki, A., Bensch, K. G. and Fleming, J. E. (1991). "Expression of bovine superoxide dismutase in *Drosophila melanogaster* augments resistance of oxidative stress." Mol Cell Biol **11**(2): 632-40.

Rincon, M., Muzumdar, R., Atzmon, G. and Barzilai, N. (2004). "The paradox of the insulin/IGF-1 signaling pathway in longevity." Mech Ageing Dev **125**(6): 397-403.

Ristoff, E. and Larsson, A. (2002). "Oxidative stress in inborn errors of metabolism: Lessons from glutathione deficiency." J Inherit Metab Dis **25**(3): 223-6.

Rogina, B., Reenan, R. A., Nilsen, S. P. and Helfand, S. L. (2000). "Extended life-span conferred by cotransporter gene mutations in *Drosophila*." Science **290**(5499): 2137-40.

Rogina, B., Helfand, S. L. and Frankel, S. (2002). "Longevity regulation by *Drosophila* Rpd3 deacetylase and caloric restriction." Science **298**(5599): 1745.

Rogina, B. and Helfand, S. L. (2004). "Sir2 mediates longevity in the fly through a pathway related to calorie restriction." Proc Natl Acad Sci U S A **101**(45): 15998-6003.

Roman, G., Endo, K., Zong, L. and Davis, R. L. (2001). "P{switch}, a system for spatial and temporal control of gene expression in *Drosophila melanogaster*." PNAS **98**(22): 12602-12607.

Rong, Y. S. and Golic, K. G. (2000). "Gene targeting by homologous recombination in *Drosophila*." Science **288**(5473): 2013-8.

Rose, M. R. and Charlesworth, B. (1981). "Genetics of life history in *Drosophila melanogaster*. I. Sib analysis of adult females." Genetics **97**(1): 173-86.

Rose, M. R. (1999). "Genetics of aging in *Drosophila*." Exp Gerontol **34**(5): 577-85.

Rose, M. R. and Long, A. D. (2002). "Ageing: The many-headed monster." Curr Biol **12**(9): R311-2.

Ruan, H., Tang, X. D., Chen, M. L., Joiner, M. L., Sun, G., Brot, N., Weissbach, H., Heinemann, S. H., Iverson, L., Wu, C. F. and Hoshi, T. (2002). "High-quality life extension by the enzyme peptide methionine sulfoxide reductase." Proc Natl Acad Sci U S A **99**(5): 2748-53.

Ruiz-Dubreuil, G., Burnet, B., Connolly, K. and Furness, P. (1996). "Larval foraging behaviour and competition in *Drosophila melanogaster*." Heredity **76 (Pt 1)**: 55-64.

Ryder, E. and Russell, S. (2003). "Transposable elements as tools for genomics and genetics in *Drosophila*." Brief Funct Genomic Proteomic **2**(1): 57-71.

Sampayo, J. N., Gill, M. S. and Lithgow, G. J. (2003). "Oxidative stress and aging-the use of superoxide dismutase/catalase mimetics to extend lifespan." Biochem Soc Trans **31**(Pt 6): 1305-7.

Sampayo, J. N. and Lithgow, G. J. (2004). "S.W.A.T.-SOD weapons and tactics." Sci Aging Knowledge Environ **2004**(25): pe27.

Sanz, A., Caro, P., Ayala, V., Portero-Otin, M., Pamplona, R. and Barja, G. (2006). "Methionine restriction decreases mitochondrial oxygen radical generation and leak as well as oxidative damage to mitochondrial DNA and proteins." Faseb J **20**(8): 1064-73.

Saunders, R. D. and McLellan, L. I. (2000). "Molecular cloning of *Drosophila* gamma-glutamylcysteine synthetase by functional complementation of a yeast mutant." FEBS Lett **467**(2-3): 337-40.

Schafer, F. Q. and Buettner, G. R. (2001). "Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple." Free Radic Biol Med **30**(11): 1191-212.

Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., Dixon, J. E. and Zipursky, S. L. (2000). "*Drosophila* dSCAM is an axon guidance receptor exhibiting extraordinary molecular diversity." Cell **101**(6): 671-684.

Schneider, I. (1972). "Cell lines derived from late embryonic stages of *Drosophila melanogaster*." J Embryol Exp Morphol **27**(2): 353-65.

Schriner, S. E., Linford, N. J., Martin, G. M., Treuting, P., Ogburn, C. E., Emond, M., Coskun, P. E., Ladiges, W., Wolf, N., Van Remmen, H., Wallace, D. C. and Rabinovitch, P. S. (2005). "Extension of murine life span by overexpression of catalase targeted to mitochondria." Science **308**(5730): 1909-1911.

Schulz, J. B., Lindenau, J., Seyfried, J. and Dichgans, J. (2000). "Glutathione, oxidative stress and neurodegeneration." Eur J Biochem **267**(16): 4904-11.

Seelig, G. F., Simondsen, R. P. and Meister, A. (1984). "Reversible dissociation of gamma-glutamylcysteine synthetase into two subunits." J Biol Chem **259**(15): 9345-7.

Seo, Y. J., Lee, J. W., Lee, E. H., Lee, H. K., Kim, H. W. and Kim, Y. H. (2004). "Role of glutathione in the adaptive tolerance to H_2O_2." Free Radic Biol Med **37**(8): 1272-81.

Seong, K. H., Matsuo, T., Fuyama, Y. and Aigaki, T. (2001). "Neural-specific overexpression of *Drosophila* plenty of SH3s (dposh) extends the longevity of adult flies."

Biogerontology **2**(4): 271-81.
Seroude, L. (2002). "Gal4 drivers expression in the whole adult fly." Genesis **34**(1-2): 34-8.
Seroude, L., Brummel, T., Kapahi, P. and Benzer, S. (2002). "Spatio-temporal analysis of gene expression during aging in *Drosophila melanogaster*." Aging Cell **1**(1): 47-56.
Seto, N. O., Hayashi, S. and Tener, G. M. (1990). "Overexpression of Cu-Zn superoxide dismutase in *Drosophila* does not affect life-span." Proc Natl Acad Sci U S A **87**(11): 4270-4.
Sgro, C. M. and Partridge, L. (1999). "A delayed wave of death from reproduction in *Drosophila*." Science **286**(5449): 2521-4.
Shanks, E. J. (2004). The regulation of thiol antioxidants in *Drosophila melanogaster* Biomedical Research Centre. Dundee, University of Dundee.
Shen, D., Dalton, T. P., Nebert, D. W. and Shertzer, H. G. (2005). "Glutathione redox state regulates mitochondrial reactive oxygen production." J Biol Chem **280**(27): 25305-12.
Shi, M. M., Kugelman, A., Iwamoto, T., Tian, L. and Forman, H. J. (1994). "Quinone-induced oxidative stress elevates glutathione and induces gamma-glutamylcysteine synthetase activity in rat lung epithelial L2 cells." J Biol Chem **269**(42): 26512-7.
Shi, Z. Z., Carter, B. Z., Habib, G. M., He, X., Sazer, S., Lebovitz, R. M. and Lieberman, M. W. (1996). "A single mouse glutathione synthetase gene encodes six mRNAs with different 5' ends." Arch Biochem Biophys **331**(2): 215-24.
Shi, Z. Z., Osei-Frimpong, J., Kala, G., Kala, S. V., Barrios, R. J., Habib, G. M., Lukin, D. J., Danney, C. M., Matzuk, M. M. and Lieberman, M. W. (2000). "Glutathione synthesis is essential for mouse development but not for cell growth in culture." Proc Natl Acad Sci U S A **97**(10): 5101-6.
Simon, A. F., Shih, C., Mack, A. and Benzer, S. (2003). "Steroid control of longevity in *Drosophila melanogaster*." Science **299**(5611): 1407-10.
Simon, A. F., Liang, D. T. and Krantz, D. E. (2006). "Differential decline in behavioral performance of *Drosophila melanogaster* with age." Mech Ageing Dev **127**(7): 647-51.
Singh, N. P., McCoy, M. T., Tice, R. R. and Schneider, E. L. (1988). "A simple technique for quantitation of low levels of DNA damage in individual cells." Exp Cell Res **175**(1): 184-91.
Sledz, C. A. and Williams, B. R. (2005). "RNA interference in biology and disease." Blood **106**(3): 787-94.
Sohal, R. S., Sohal, B. H. and Orr, W. C. (1995). "Mitochondrial superoxide and hydrogen peroxide generation, protein oxidative damage, and longevity in different species of flies." Free Radic Biol Med **19**(4): 499-504.
Sohal, R. S. and Weindruch, R. (1996). "Oxidative stress, caloric restriction, and aging." Science **273**(5271): 59-63.
Sohal, R. S. (2002). "Role of oxidative stress and protein oxidation in the aging process." Free Radic Biol Med **33**(1): 37-44.
Sohal, R. S., Mockett, R. J. and Orr, W. C. (2002). "Mechanisms of aging: An appraisal of the oxidative stress hypothesis." Free Radic Biol Med **33**(5): 575-86.
Soltaninassab, S. R., Sekhar, K. R., Meredith, M. J. and Freeman, M. L. (2000). "Multi-faceted regulation of gamma-glutamylcysteine synthetase." J Cell Physiol **182**(2): 163-70.
Spencer, C. C., Howell, C. E., Wright, A. R. and Promislow, D. E. (2003). "Testing an 'aging gene' in long-lived *Drosophila* strains: Increased longevity depends on sex and genetic background." Aging Cell **2**(2): 123-30.
Spradling, A. C. and Rubin, G. M. (1982). "Transposition of cloned p elements into *Drosophila* germ line chromosomes." Science **218**(4570): 341-7.

Srivastava, S. K., Awasthi, Y. C. and Beutler, E. (1974). "Useful agents for the study of glutathione metabolism in erythroyctes. Organic hydroperoxides." Biochem J **139**(2): 289-95.

Stadtman, E. R. (2002). "Importance of individuality in oxidative stress and aging." Free Radic Biol Med **33**(5): 597-604.

Stadtman, E. R. and Levine, R. L. (2003). "Free radical-mediated oxidation of free amino acids and amino acid residues in proteins." Amino Acids **25**(3-4): 207-18.

Sun, J. and Tower, J. (1999). "Flp recombinase-mediated induction of Cu/Zn-superoxide dismutase transgene expression can extend the life span of adult *Drosophila melanogaster* flies." Mol Cell Biol **19**(1): 216-28.

Sun, J., Folk, D., Bradley, T. J. and Tower, J. (2002). "Induced overexpression of mitochondrial Mn-superoxide dismutase extends the life span of adult *Drosophila melanogaster*." Genetics **161**(2): 661-72.

Tatar, M., Kopelman, A., Epstein, D., Tu, M. P., Yin, C. M. and Garofalo, R. S. (2001). "A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function." Science **292**(5514): 107-10.

Tatar, M., Bartke, A. and Antebi, A. (2003). "The endocrine regulation of aging by insulin-like signals." Science **299**(5611): 1346-51.

Tatusova, T. A. and Madden, T. L. (1999). "Blast 2 sequences, a new tool for comparing protein and nucleotide sequences." FEMS Microbiol Lett **174**(2): 247-50.

Thiruchelvam, M., Prokopenko, O., Cory-Slechta, D. A., Richfield, E. K., Buckley, B. and Mirochnitchenko, O. (2005). "Overexpression of superoxide dismutase or glutathione peroxidase protects against the paraquat + maneb-induced parkinson disease phenotype." J Biol Chem **280**(23): 22530-9.

Thornton, K. and Long, M. (2002). "Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome." Mol Biol Evol **19**(6): 918-25.

Tice, R. R., Agurell, E., Anderson, D., Burlinson, B., Hartmann, A., Kobayashi, H., Miyamae, Y., Rojas, E., Ryu, J. C. and Sasaki, Y. F. (2000). "Single cell gel/comet assay: Guidelines for in vitro and in vivo genetic toxicology testing." Environ Mol Mutagen **35**(3): 206-21.

Tietze, F. (1969). "Enzymic method for quantitative determination of nanogram amounts of total and oxidized glutathione: Applications to mammalian blood and other tissues." Anal Biochem **27**(3): 502-522.

Tower, J. (2000). "Transgenic methods for increasing *Drosophila* life span." Mech Ageing Dev **118**(1-2): 1-14.

Tower, J. (2002). "Setting a trap for aging-related genes in *Drosophila*." Sci Aging Knowledge Environ **2002**(39): pe15.

Townsend, D. M., Tew, K. D. and Tapiero, H. (2003). "The importance of glutathione in human disease." Biomed Pharmacother **57**(3-4): 145-55.

Ullmann, P., Gondet, L., Potier, S. and Bach, T. J. (1996). "Cloning of *Arabidopsis thaliana* glutathione synthetase (GSH2) by functional complementation of a yeast GSH2 mutant." Eur J Biochem **236**(2): 662-9.

Valenzuela, R. K., Forbes, S. N., Keim, P. and Service, P. M. (2004). "Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. II. Response to selection." Genetics **168**(1): 313-24.

Valverde, M., Rojas, E., Kala, S. V., Kala, G. and Lieberman, M. W. (2006). "Survival and cell death in cells constitutively unable to synthesize glutathione." Mutat Res **594**(1-2): 172-80.

Van Remmen, H., Qi, W., Sabia, M., Freeman, G., Estlack, L., Yang, H., Mao Guo, Z., Huang, T. T., Strong, R., Lee, S., Epstein, C. J. and Richardson, A. (2004). "Multiple

deficiencies in antioxidant enzymes in mice result in a compound increase in sensitivity to oxidative stress." Free Radic Biol Med **36**(12): 1625-34.
Van Remmen, H. and Richardson, A. (2001). "Oxidative damage to mitochondria and aging." Exp Gerontol **36**(7): 957-68.
Vina, J., Sastre, J., Pallardo, F. and Borras, C. (2003). "Mitochondrial theory of aging: Importance to explain why females live longer than males." Antioxid Redox Signal **5**(5): 549-56.
Walker, D. W., Muffat, J., Rundel, C. and Benzer, S. (2006). "Overexpression of a *Drosophila* homolog of apolipoprotein d leads to increased stress resistance and extended lifespan." Curr Biol **16**(7): 674-9.
Wallace, D. C. (1999). "Mitochondrial diseases in man and mouse." Science **283**(5407): 1482-8.
Wallenfang, M. R., Nayak, R. and Dinardo, S. (2006). "Dynamics of the male germline stem cell population during aging of *Drosophila melanogaster*." Aging Cell. **5**(4): 297-304
Walsh, J. B. (1995). "How often do duplicated genes evolve new functions?" Genetics **139**(1): 421-8.
Wang, C. L. and Oliver, D. J. (1996). "Cloning of the cdna and genomic clones for glutathione synthetase from *Arabidopsis thaliana* and complementation of a Gsh2 mutant in fission yeast." Plant Mol Biol **31**(6): 1093-104.
Wang, C. L. and Oliver, D. J. (1997). "Glutathione synthetase: Similarities of the proteins from *Schizosaccharomyces pombe* and *Arabidopsis thaliana*." Biochem J **326 (Pt 2)**: 563-6.
Wang, H. D., Kazemi-Esfarjani, P. and Benzer, S. (2004). "Multiple-stress analysis for isolation of *Drosophila* longevity genes." Proc Natl Acad Sci U S A **101**(34): 12610-5.
Wang, J. W., Humphreys, J. M., Phillips, J. P., Hilliker, A. J. and Wu, C. F. (2000). "A novel leg-shaking *Drosophila* mutant defective in a voltage-gated K(+)current and hypersensitive to reactive oxygen species." J Neurosci **20**(16): 5958-64.
Wang, M. C., Bohmann, D. and Jasper, H. (2003). "JNK signaling confers tolerance to oxidative stress and extends lifespan in *Drosophila*." Dev Cell **5**(5): 811-6.
Wang, M. C., Bohmann, D. and Jasper, H. (2005). "JNK extends life span and limits growth by antagonizing cellular and organism-wide responses to insulin signaling." Cell **121**(1): 115-25.
Watson, W. H. and Jones, D. P. (2003). "Oxidation of nuclear thioredoxin during oxidative stress." FEBS Lett **543**(1-3): 144-7.
Webb, G. C., Vaska, V. L., Gali, R. R., Ford, J. H. and Board, P. G. (1995). "The gene encoding human glutathione synthetase (GSS) maps to the long arm of chromosome 20 at band 11.2." Genomics **30**(3): 617-9.
Wei, Y. H., Lu, C. Y., Wei, C. Y., Ma, Y. S. and Lee, H. C. (2001). "Oxidative stress in human aging and mitochondrial disease-consequences of defective mitochondrial respiration and impaired antioxidant enzyme system." Chin J Physiol **44**(1): 1-11.
Whitbread, L., Gali, R. R. and Board, P. G. (1998). "The structure of the human glutathione synthetase gene." Chemico-Biological Interactions **111-112**: 35-40.
Wilson, R., Morgan, T. and Mackay, T. F. C. (2006). "High resolution mapping of quantitative trait loci affecting increased life span in *Drosophila melanogaster*." Genetics: genetics.105.055111.
Wong, C. M., Zhou, Y., Ng, R. W., Kung Hf, H. F. and Jin, D. Y. (2002). "Cooperation of yeast peroxiredoxins Tsa1P and Tsa2P in the cellular defense against oxidative and nitrosative stress." J Biol Chem **277**(7): 5385-94.
Wong, K. K., Maser, R. S., Bachoo, R. M., Menon, J., Carrasco, D. R., Gu, Y., Alt, F. W. and DePinho, R. A. (2003). "Telomere dysfunction and Atm deficiency compromises

organ homeostasis and accelerates ageing." Nature **421**(6923): 643-8.
Worby, C. A., Simonson-Leff, N. and Dixon, J. E. (2001). "RNA interference of gene expression (RNAi) in cultured *Drosophila* cells." Sci. STKE **2001**(95): pl1-.
Yanase, S., Yasuda, K. and Ishii, N. (2002). "Adaptive responses to oxidative damage in three mutants of *Caenorhabditis elegans* (age-1, mev-1 and daf-16) that affect life span." Mech Ageing Dev **123**(12): 1579-87.
Yang, H., Zeng, Y., Lee, T. D., Yang, Y., Ou, X., Chen, L., Haque, M., Rippe, R. and Lu, S. C. (2002). "Role of AP-1 in the coordinate induction of rat glutamate-cysteine ligase and glutathione synthetase by tert-butylhydroquinone." J Biol Chem **277**(38): 35232-9.
Yang, H., Magilnick, N., Lee, C., Kalmaz, D., Ou, X., Chan, J. Y. and Lu, S. C. (2005a). "Nrf1 and Nrf2 regulate rat glutamate-cysteine ligase catalytic subunit transcription indirectly via Nf-kappa-b and AP-1." Mol Cell Biol **25**(14): 5933-46.
Yang, H., Magilnick, N., Ou, X. and Lu, S. C. (2005b). "Tumour necrosis factor alpha induces co-ordinated activation of rat GSH synthetic enzymes via nuclear factor kappa-b and activator protein-1." Biochem J **391**(Pt 2): 399-408.
Yang, X., Wu, X., Choi, Y. E., Kern, J. C. and Kehrer, J. P. (2004). "Effect of acrolein and glutathione depleting agents on thioredoxin." Toxicology **204**(2-3): 209-18.
Yang, Y., Dieter, M. Z., Chen, Y., Shertzer, H. G., Nebert, D. W. and Dalton, T. P. (2002). "Initial characterization of the glutamate-cysteine ligase modifier subunit gclm(-/-) knockout mouse. Novel model system for a severely compromised oxidative stress response." J Biol Chem **277**(51): 49446-52.
Yang, Y., Gehrke, S., Haque, M. E., Imai, Y., Kosek, J., Yang, L., Beal, M. F., Nishimura, I., Wakamatsu, K., Ito, S., Takahashi, R. and Lu, B. (2005). "Inactivation of *Drosophila* DJ-1 leads to impairments of oxidative stress response and phosphatidylinositol 3-kinase/akt signaling." Proc Natl Acad Sci U S A **102**(38): 13670-5.
Yu, B. P., Masoro, E. J., Murata, I., Bertrand, H. A. and Lynd, F. T. (1982). "Life span study of SPF fischer 344 male rats fed ad libitum or restricted diets: Longevity, growth, lean body mass and disease." J Gerontol **37**(2): 130-41.
Zhang, J. (2003). "Evolution by gene duplication: An update." TRENDS in Ecology and Evolution **18**(6): 292-298.
Zheng, J., Edelman, S. W., Tharmarajah, G., Walker, D. W., Pletcher, S. D. and Seroude, L. (2005). "Differential patterns of apoptosis in response to aging in *Drosophila*." PNAS **102**(34): 12083-12088.
Zhu, Y., Carvey, P. M. and Ling, Z. (2006). "Age-related changes in glutathione and glutathione-related enzymes in rat brain." Brain Res **1090**(1): 35-44.
Zou, S., Meadows, S., Sharp, L., Jan, L. Y. and Jan, Y. N. (2000). "Genome-wide study of aging and oxidative stress response in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **97**(25): 13726-31.